ANALYSIS

Edited by Christopher Kirwan

15

Volume 46

1986

ISSN 0003-2638

BASIL BLACKWELL · OXFORD

INDEX

		PAC
ALEXANDER, PETER, *et al.*	Open letter to the international philosophical community	16
BASSHAM, GREGORY	Ehring's theory of causal asymmetry	2
BUCKNER, D. K.	Transparently false: reply to Hardin	8
BUCKNER, D. K. *and* SMITH, PETER	Quotation and the Liar Paradox	6
CLARK, ANDY	Superman, the image	22
COOK, MONTE	Indeterminacy of identity	17
COSTA, MICHAEL J.	Acting intentionally and minimal abilities	14
COSTA, MICHAEL J.	Hume's argument for the temporal priority of cause over effect	8
CRISP, ROGER	The avoidance of the problem of evil: a reply to McGrath	16
DALE, A. J.	'If A and B, then A'	8
DALE, A. J.	Thing statements and appearance statements	2
DENISE, THEODORE C.	Redundancy and *INUS* conditionality	12
DUFF, ANTONY	Pascal's wager and infinite utilities	10
EHRING, DOUGLAS	Closed causal loops, single causes, and asymmetry	3
FISCHER, JOHN MARTIN	Pike's Ockhamism	5
GOLDSTEIN, LAURENCE	Epimenides and Curry	11
GOLDSTEIN, LAURENCE	False stipulation and semantical paradox	19
GRIM, PATRICK	On sets and worlds: a reply to Menzel	18
HARRISON, JONATHAN	Correction	11
HOLLIS, MARTIN	More paradoxical epistemics	21
JACOBSON, ANNE JAAP	Causality and the supposed counterfactual conditional in Hume's *Enquiry*	13
JACOBSON, ANNE JAAP	Correction	22
JENNINGS, RICHARD C.	Tarski – an ambiguity	20
KINGHAN, MICHAEL	A paradox derailed: reply to Hollis	2
LEAMAN, OLIVER	Harris on non-vicious choice	10
LEMKE, LORY	Kvanvig and Swain on the basing relation	1

Linsky, Bernard	Factives, blindspots and some paradoxes	10
Lowe, E. J.	On a supposed temporal/modal parallel	195
Lowe, E. J.	Noonan on naming and predicating	159
Lowe, E. J.	Reply to Dale	83
Lowe, E. J.	Reply to Over	200
McGrath, P. J.	Evil and the existence of a finite God	63
Melchert, Norman	Metaphysical realism and history	36
Mellor, D. H.	Tense's tenseless truth conditions	167
Menzel, Christopher	On set theoretic possible worlds	68
Montmarquet, James	Prosthesis and pre-emption	147
Montmarquet, James	The voluntariness of belief	49
Moore, A. W.	How significant is the use/mention distinction?	173
Naylor, Margery Bedford	A note on David Lewis's realism about possible worlds	28
Newton, Natika	Churchland on direct introspection of brain states	97
Noonan, Harold W.	AE or EA?	87
Noonan, Harold W.	Relative identity: a reconsideration	6
Noonan, Harold W.	Reply to Garrett	205
Noonan, Harold W.	Reply to Lowe	218
Olin, Doris	On a paradoxical train of thought	18
Over, D. E.	Is there a temporal slippery slope paradox?	197
Over, D. E.	On a temporal slippery slope paradox	15
Over, D. E.	Reply to Lowe	201
Peacocke, Christopher	Intention and practical reasoning: a reply to Donald Davidson	45
Pettit, Phillip	*A priori* principles and action-explanation	39
Priest, Graham	Tense and truth conditions	162
Rein, Andrew	Repeatable miracles?	109
Sanford, David H.	The possibility of transparent white	212
Schlesinger, George N.	On the limits of science	24
Shand, John	Grayling, Feyerabend and the constancy of sense	211
Slater, B. H.	Prior's analytic	76
Smith, Peter	Metaphysical realism and historical interpretation	157
Smith, Peter	(*see* Buckner, D. K.)	

SORENSEN, ROY A.	Nozick, justice, and the sorites	102
SPINKS, GRAHAM	Noonan on temporal parts	215
THALBERG, IRVING	The immateriality of 'abstract objects' and the mental	93
TIENSON, JOHN	An observation on common names and proper names	73
WAGNER, STEVEN J.	Quine's holism	1
WATERS, C. KENNETH	Why Chisholm's analysis of justification won't do	134
WHITE, ALAN R.	Ways of speaking of imagination	152
WILLIAMS, S. G.	On the logical possibility of time without change	122
WILLIAMSON, TIMOTHY	The contingent *a priori*: has it anything to do with indexicals?	113
ZOCH, L. N.	Remorse and regret: a reply to Phillips and Price	54

103
Cde-H02973-15- An 13 P004J08

l. 46 No. 1 (New Series No. 209) January 1986

ANALYSIS

Edited by

CHRISTOPHER KIRWAN

CONTENTS

Quine's holism	STEVEN J. WAGNER
Relative identity: a reconsideration	HAROLD W. NOONAN
Factives, blindspots and some paradoxes	BERNARD LINSKY
On a temporal slippery slope paradox	D. E. OVER
On a paradoxical train of thought	DORIS OLIN
A paradox derailed: reply to Hollis	MICHAEL KINGHAN
On the limits of science	GEORGE N. SCHLESINGER
Thing statements and appearance statements	A. J. DALE
A note on David Lewis's realism about possible worlds	MARGERY BEDFORD NAYLOR
Ehring's theory of causal asymmetry	GREGORY BASSHAM
Closed causal loops, single causes, and asymmetry	DOUGLAS EHRING
Metaphysical realism and history	NORMAN MELCHERT
A priori principles and action-explanation	PHILIP PETTIT
Intention and practical reasoning: a reply to Donald Davidson	CHRISTOPHER PEACOCKE
The voluntariness of belief	JAMES MONTMARQUET
Remorse and regret: a reply to Phillips and Price	L. N. ZOCH
Pike's Ockhamism	JOHN MARTIN FISCHER
Evil and the existence of a finite God	P. J. MCGRATH

BASIL BLACKWELL PUBLISHER · OXFORD

ISSN 0003-2638

ANALYSIS 46.1 JANUARY 1986

105
Ar 13

QUINE'S HOLISM

By STEVEN J. WAGNER

TWO features of Quine's approach to meaning stand out. First, it is verificationist: the kind of meaning science recognizes is a matter of sensory evidence. Second, it is holist: only large theories, not statements (except for some very long ones), have meaning. These features are linked, of course. Quine denies meaningfulness for statements because he thinks they lack verification conditions. Meaningful bodies of language must be large enough to admit confirmation (and infirmation). Thus a doctrine about confirmation lies at the foundation of Quine's view of meaning; and it is the subject of this note.

Holism about confirmation (c-holism) plays a major role in 'Two Dogmas of Empiricism' [1]. Recall that while Quine there dismisses numerous arguments for an analytic-synthetic distinction, he does not reject the distinction itself on that ground alone. His positive reason derives from c-holism: since statements cannot have meanings (i.e., for Quine, verification meanings), they cannot, *a fortiori*, be analytic. But his case for c-holism could hardly bear such an important weight, for it involves a plain fallacy. Quine begins with a suggestive metaphor:

> (a) No statement, taken in isolation from its fellows, admits confirmation at all. Our statements face the tribunal of sense experience only as a corporate body.

From this he moves to

> (b) It is not significant, in general, to speak of the confirmation of a single statement.

But the opposite follows. If s admits of confirmation when *not* taken in isolation from its fellows, it necessarily admits of confirmation.

Quine might well reply that the verification conditions of s in theory T are *trivial* in the sense that s shares these with every other statement in T. If, say, we derive an observation statement o from T and see that $\sim o$ obtains, then T is infirmed, and so, therefore, are all the statements that comprise it. But all are equally, indiscriminately infirmed. If we regard our test as more relevant to some than to others, we are making a conventional choice. 'Two Dogmas,' however, contains no argument for this – *prima facie* implausible – reply.

The difficulty here is an ambiguity in claims like (a), hence in the notion of holism. A weak reading of (a) makes verification conditions dependent on surrounding theory. Thus, whether we should accept *h* on the basis of *o* may depend on how simple or coherent that would render our total theory. This is the thesis suggested by much of 'Two Dogmas' and by the passages in Carnap on which Quine drew. It is a major insight; but it does not make single hypotheses unverifiable. A given observation may still confirm *h* to a certain degree without equally confirming sentences in the surrounding theory. Similarly, data relevant to the containing doctrine may not bear on *h*. Distinct sentences of theory may thus have distinct sets of degrees to which various observations confirm them, even though the verification of each involves its connections to other sentences. So we could define the meaning of *s* as the set of all pairs (O, D) such that O would confirm *s* to degree D in the context of *T*: this is holistic in its relativization to a theory, yet it continues to allow testing of sentences.[1] The stronger reading of (a) rejects this picture, claiming that 'testing *h*' can only mean testing a containing *T* that is evaluated as a unit: given two sentences *h* and *h'* in *T*, there is no observation more relevant to one than to the other. This is the intended meaning of (b), which I am calling c-holism. It is what the attack on analyticity in 'Two Dogmas' requires. Evidently it does not follow from the weaker reading of (a). Quine's problem, then, is to show why the epistemology of confirmation should lead us to c-holism, not just to an intermediate position that allows the holistic confirmation of individual hypotheses. For the attempt to bridge this gap we must turn to the first chapter of [2].

There Quine remarks that when a biologist devises an experiment to test an hypothesis *h*, she has antecedently fixed upon *h* as more tentative and suspect than the rest of her theory (p. 7). This is true, but *by itself* it begs the question. At issue is why *h* should have been selected. If that was arbitrary, Quine is right, but if the data were in fact, independently of our choices, evidence for or against *h* in particular, then the biologist's procedure simply reflected objective confirmation relations. The argument with which Quine continues in the text does not tackle this issue and quite fails to advance his case. He notes that even when we try to question just one hypothesis, we simultaneously question any that (in some sense) imply it, so that no hypothesis is ever tested alone. Although this is again correct, it is irrelevant. We might carelessly suppose that we can test exactly one hypothesis in an experiment, but the question is whether intuitively distinct hypothesis have different verification conditions. And we can agree with Quine's remarks while continuing to hold that since one may infirm hypotheses that imply *h* without infirming *h*, and confirm consequences of *h* but not *h* itself, *h* has its distinctive ranges of relevant observations. It is one thing to argue that whenever *h* is tested, some further body of theory is tested in the bargain, and another to conclude (as Quine needs to) that there is a theory *T* such that whenever *h* is tested, *T* is tested. One would not wish to accuse Quine of this fallacious transposition. Yet his meaning otherwise remains unclear. We do better to turn to the immediately preceding paragraph of [2], which I take to be Quine's most considered attempt to bring out what was going on in 'Two Dogmas':

---

[1] A different notion results if we let the theoretical context vary, identifying the meaning of *s* with the set of all triples (*T*, O, D) where *T* is *any* theory containing *s*, and O and D are as before. This notion, however, is useless to a verificationist because the corresponding synonymy relation is too strong – it coincides with logical equivalence. Anyway, both notions run into the problems about embedding and intertheoretic translation noted below, so Quine could use neither.

> The hopelessness of distributing empirical information generally over separate sentences, or even over fairly large bundles of sentences, is in some sense widely recognized, if only by implication ... It will be widely agreed that our theory of nature is underdetermined by ... all "possible" observations. This means that there can be a set H of hypotheses, and an alternative set H' incompatible with H, and it can happen that when our total theory T is changed to the extent of putting H for H' in it, the resulting theory T' still fits all possible observations just as well as T did. Evidently then H and H' convey the same empirical information, as far as empirical information can be apportioned to H and H' at all; but still they are incompatible. This reflection should scotch any general notion of ... [the] empirical [meaningfulness] of sentences (pp. 6–7).

Thus the attempt to assign empirical meanings is supposed to lead to the absurd assignment of empirical synonomy to contradictory hypotheses. But this argument is curious.

(1) Why must H and H' convey the same empirical information? We should expect Quine to hold that the empirical content of a part of a theory T is no different from that of T itself (the empirical equivalence of H and H' then follows from that of T and T'), but the interest of our quotation depends on its being an *argument* for this view. Perhaps Quine assumes that if T and T' convey the same empirical information (since both fit all possible observations equally well), and if T' is obtained by the substitution of H' for H, then H and H' must convey the same information. But to take an analogous example, it is false that if s and s' are *logically* equivalent, where s' is obtained from s by substituting p' for p, then p and p' are logically equivalent. Since the inductive case ought to be trickier than the deductive one, Quine's assumption would be puzzling. But I cannot otherwise explain his transition to the empirical equivalence of H and H'.

(2) Just why can't two incompatible hypotheses have the same empirical content? It is on this basis, after all, that Quine rejects the assignment of empirical content to H and H' without further discussion, but I see no contradiction. A familiar idea from the philosophy of science is that one observation can, in the contexts

of different background theories, count for and against a given hypothesis, and the same phenomenon is at work in Quine's story. Suppose that *H* contains hypotheses about demonic possession, whereas H' denies these in favour of talk of brain disturbances. It seems that the resulting disagreement between the total containing theories *T* and T' might not be resoluble by any observation anywhere in the universe. Even criteria of simplicity and coherence might not yield a decision. In this situation *H* and H' could, if properly chosen – if carved out of *T* and T' so as to include, along with the materials about demons and brains, suitable auxiliary hypotheses – be empirically equivalent in spite of their incompatibility. Whatever would confirm *H* in the context of *T* might confirm H' in the context of T', and vice versa.

Of course I have not shown that this really is possible. I may also appear to have begged questions against Quine by assuming determinate, distinctive contents for *H* and H', and by treating the vocabularies of *T* and T' as intertranslatable ("commensurable"). But such charges would miss the point. Quine is trying to *demonstrate* an *inconsistency* in a standard picture of confirmation. Clearly he has raised no difficulty at all.

Why, then, does Quine resort to these confused arguments against sentence meaning? One answer is that since c-holism is false – if we allow that – arguments for it are likely to be weak. But this only drives the question back to why Quine should maintain c-holism. Also, it leaves unexplained the particular errors we have found.

These questions gain force when one notes that the attack on analyticity needs nothing as strong as c-holism. One can hold that no statements are immune to revision without making the extreme move of denying them any meaning. Of course, that would leave the Quine of 'Two Dogmas' without a ground for his critique of analyticity, but that could be remedied. *Word & Object* (M.I.T. 1960) tries to develop a philosophy of language that simply does without analyticity and related traditional notions; if successful, that would suffice for their abandonment. Indeed, Quine would need this sort of argument anyway. As long as one recognizes any meaningful units of language at all, the question of their analyticity can be raised. One may, for example, ask whether any comprehensive theories are immune to revision. A notion of analyticity confined to such theories might be less interesting – it might not invite the same applications – but it would still be analyticity and would have to be ruled out on grounds other than c-holism.

Note also that if the argument of [2] were sound, it would buy trouble for Quine. Its basic idea is the fact about confirmation already suggested in 'Two Dogmas,' which we may call the embedding phenomenon: the relevance of sensory evidence to *s* depends on the position of *s* in surrounding theory. (This is what we turned against Quine in (2) above.) Quine exploits this as follows. Given the assumption that a unit of language *U* conveys empirical information, he observes that *U* could be embedded in a larger unit in a way that would effect its empirical equivalence with a likewise embedded, incompatible *U'*. That is supposed to be absurd. Now this reasoning does not depend on the specific character of *U*, as long as *U* admits embedding. But everything can be embedded. The verification conditions even of a theory *T* large enough to have, in any sense, empirical content of its own will still vary with *T*'s background. Where *T* is the whole of science (assuming that to be well defined), it is indeed trivially true that *T* is not embedded in anything on which its verification conditions depend, but they do depend on how that particular whole is embedded. If it is not embedded at all, *T* has one set of such conditions, and it will have others if it is embedded in other bodies of information. Since the empirical content of *T* can change with the addition of a single statement *s*, we will have to distinguish the meaning of *T* when *T* was the whole of science from its meaning in the new $T + s$. This is not necessarily absurd, but it opens the door to Quine's argument. Suitable embeddings could bring about the empirical equivalence of two incompatible wholes of science, thus disqualifying even such wholes as units of empirical significance. (Note that Quine's *T* and *T'* should themselves be incompatible, so it should be just as absurd to assign empirical meanings to them as to *H* and *H'*. [3] voices doubts about the possibility of empirically equivalent, *incompatible* theories. My concern is Quine's philosophy before then; the shift brings problems of its own.)

We have not only seen that c-holism fails to serve at least one of its main purposes, but also that Quine's argument for it is positively deleterious. So why not drop it? Perhaps it is a necessary (for Quine) concession from epistemology to the philosophy of language. Quine the epistemologist saw how embedding determined verification conditions. But Quine also insisted on verificationism, and within it, c-holism is the closest one can come to describing this phenomenon. Verificationism makes the identity of a statement depend on its verification conditions: something with different empirical content is simply a different statement. Recognizing embedding, however, involves reidentifying statements across changes of theory, since we want to say that the verification conditions of a statement will vary with its theoretical context.[2] So it seems that the best the verificationist can do is to deny that individual statements are verifiable at all. He will sense that the embedding problem makes a verificationist account of the meanings of statements unacceptable, but will also remain committed to identifying meanings with verification conditions if such there are. C-holism is the only way out. The epistemological loss, of course, is that embedding can no longer be described.

---

[2] Similarly for the weak reading of (a). Saying that an observation unequally infirms two sentences in *T* means, roughly, that we may be justified in moving to a new *T'*, containing just one of them. But this assumes reidentifiability across theories.

These reflections do illuminate Quine's reasoning, as well as the thesis of c-holism itself. One can see the failed attempts to move from a view of evidence that recognizes embedding – holism in one sense – to a rejection of distinctive verification conditions for statements – holism in the other. Quine would have done better to combine verificationism with the recognition of a second sense of meaning tied not to evidence but, presumably, to truth conditions. Such a ('two factors') position can do justice both to the fact of embedding and to the intuitions behind verificationism, as recent authors have noted. But that move lies beyond the framework of 'Two Dogmas' – even though it would hardly have been possible without Quine's contributions.

University of Illinois,
810 South Wright Street,
Urbana, Illinois 61801, U.S.A.

© STEVEN J. WAGNER 1986

REFERENCES

[1] W. V. Quine, 'Two dogmas of empiricism', *Philosophical Review* 1951, reprinted in *From a Logical Point of View*, Harvard, 2nd edition 1964.
[2] W. V. Quine, *Philosophy of Logic*, Prentice-Hall (Englewood Cliffs, N.J.: 1970).
[3] W. V. Quine, 'Empirically Equivalent Systems of the World', *Erkenntnis*, 9 (1975), 313–28.

RELATIVE IDENTITY: A RECONSIDERATION

By HAROLD W. NOONAN

IN *Objects and Identity* (The Hague, Martinus Nijhoff, 1980, hereafter *O&I*) I argued that identity was relative, in the sense that identity under a sortal concept did not entail indiscernibility, and therefore, where 'F' and 'G' were two sortal terms, x and y could be the same F but different G's.[1] I now think that neither the arguments I gave in that book, nor any arguments I have seen, in fact establish this relative identity thesis. But these arguments nonetheless establish *something*. Namely, they establish that one must maintain, for any sortal term 'F', a pair of distinctions: (a) between restricted quantification over F's and unrestricted quantification over things that are predicatively F, and (b) between names *for* F's, i.e. names associated with the criterion of identity *same F*, and names of F's, i.e. names naming things which are predicatively F.[2]

---

[1] By a sortal concept I mean, following Strawson, a concept which provides a principle for individuating and counting *particulars.* I take it to be an implication of this that things which are at some time identical under a sortal concept must then, at least partially, coincide. Thus Geach's *surman* (or my *sizeman* in *O&I*) is not in my sense a sortal concept, and rejection of the relative identity thesis as stated in the text is consistent with the thesis that one can introduce names *for* surmen (in the sense explained below) as well as names *for* men and that the predicate 'is a surman' applies univocally both to the bearers of names *for* men and to the bearers of names *for* surmen. Whether this thesis is true I do not know.

Of course, I already maintained the need for this pair of distinctions in *O&I*, where I treated it as something which immediately followed from the relative identity thesis. I still think that that was correct. But what I failed to see in that book was that if one accepted this pair of distinctions one had, in fact, no need to accept that identity was relative.

To see why this is so one needs first to appreciate the link between distinction (a) and distinction (b), which is as follows.

The restricted existentially quantified proposition 'for some F x, Ax' is true if and only if 'Ax' is true for some interpretation of 'x' as a name *for* an F, and the restricted universally quantified proposition 'for every F x, Ax' is true if and only if 'Ax' is true for every interpretation of 'x' as a name *for* an F. By contrast, the unrestricted existentially quantified proposition 'for some x, Fx and Ax' is true if and only if 'Fx and Ax' is true for some interpretation of 'x' as a name (not necessarily a name *for* an F), and the unrestricted universally quantified proposition 'for every x, if Fx then Ax' is true if and only if 'if Fx then Ax' is true for every interpretation of 'x' as a name.

Consequently, if something can be predicatively F and yet not be capable of being named by a name *for* an F, 'for some F x, Ax' is a stronger proposition than (entails but is not entailed by) 'for some x, Fx and Ax' and 'for every F x, Ax' is a weaker proposition than (does not entail, but is entailed by) 'for every x, if Fx then Ax'. This, I now believe, is what the arguments mistakenly advertised in *O&I* as arguments for relative identity in fact establish.

To see this it will be sufficient to look at the Tib/Tibbles argument (*O&I*, p. 22), which was, of course, introduced into modern discussion by Geach and, in a different version from that in my book, has been published by him as the 'Paradox of 1,001 cats' (*Reference and Generality*, last chapter). I did not, in *O&I*, put this forward as an argument for relative identity, but merely as an argument which demonstrated that it was necessary for an absolutist who would have no truc with the conception of continuants as space-time worms to embrace Wiggins' constitutive sense of 'is'

---

[2] This is Geach's terminology [*Reference and Generality*, Cornell, third edition, 1980, p. 70). My own way of making the distinction in *O&I* was to distinguish between names *of* F's, with 'of' italicised, and names of F's, but I had exactly the same distinction in mind as Geach and just expressed it differently.

(I thought this point merited an argument because I had previously shown that the putative examples of relative identity which Wiggins originally introduced the 'is' of constitution to deal with were, in fact, ones that an absolutist could handle without it). But in seeing why I was mistaken about this it will also become clear what was wrong with the arguments for relative identity which *were* put foward in *O&I.* The Tib/Tibbles argument goes like this. Take a cat called Tibbles, consider the portion of her which includes everything except her tail, and give the name Tib to that portion (Tibbles' 'puss', I called it). Since Tibbles and Tib do not occupy all of the same space at the same time, they are non-identical. But what if we amputate Tibbles' tail? Tibbles and Tib now occupy exactly the same space as each other. If Tibbles is still a cat, it is had to see by what criterion one could deny that Tib is a cat. Yet they are distinct individuals, because their histories are different. (For example, it may be true of Tibbles that she once had her tail run over, but it cannot be true of Tib: the tail was never part of her.) But there is just *one* cat in the place they now both occupy. So they cannot be distinct *cats.* They must be the same cat, even though they are distinct individuals, and so identity under the sortal concept *cat* must be a relative identity relation, i.e. a relation which does not ensure indiscernibility of its terms.

As I said in *O&I*, this argument can be met by denying that Tib is predicatively a cat when it coincides with Tibbles. But in fact it can be met even if this premiss is accepted. For the argument: 'Tib is a cat in place P, Tibbles is a cat in place P, there is just one cat in place P, therefore Tib and Tibbles are the same cat' is invalid.

'There is just one cat in P at t' means 'For some cat x, x is in P at t and for every cat y, if y is in P at t then x is the same cat as y'. Consequently, if t is the time after the amputation of Tibbles' tail, and P is the place then occupied by Tib and Tibbles, 'there is just one cat in P at t' is true (it would still be true, incidentally, if it meant 'for some cat x, x is in P at t and for every cat y, if y is in P at t then $x = y$'). And it is true *whether or not* Tib is the same cat as Tibbles; for whether or not this is so, 'Tib' is not a name *for* a cat, but a name *for* a puss, and so Tib does not fall within the domain of discourse of the restricted quantifiers 'for some cat x', 'for every cat y' which I employed in the explanation of the truth conditions of 'There is just one cat in P at t'.

Of course, this explanation of why the Tib/Tibbles argument fails assumes that something can be predicatively a cat without being capable of being named by a name *for* a cat. But my point is just that if one makes this assumption one can block the argument for relative identity the Tib/Tibbles example suggests without having to claim that Tib is only constitutively and not predicatively a cat, and consequently the arguments for relative identity I actually gave in *O&I*, which could not be met by an appeal to the 'is' of constitution, can be answered similarly.

Of course, if one accepts that Tib is predicatively a cat at t, one is not *debarred* from saying that it and Tibbles are the same cat at t. But nor, as I thought previously, is one *obliged to* say this. One has a choice. Against saying that Tib and Tibbles, when coincident, are the same cat is the fact that it is true at this time that Tib is a cat, Tibbles is a cat and Tib is not identical with Tibbles. Consequently, if one says that Tib and Tibbles are the same cat one is committed to denying that 'same cat' splits up into 'is the same and is a cat' and to accepting the relative identity thesis. In favour of saying that Tib and Tibbles, when coincident, are the same cat, is the fact that if one denies this while accepting that Tib is predicatively a cat at t then one is committed to saying that only one cat is in P at t, though Tib is a cat there, Tibbles is a cat there and Tib is not the same cat as Tibbles. But neither of these considerations is conclusive.

There is another point worth noting about the Tib/Tibbles example. It is fairly plausible to suppose that Tib is a cat before the amputation of Tibbles' tail, as well as after, and in fact is always and necessarily a cat. If so, not only is the conjunction 'Tib is a cat at t, Tib is in P at t, Tibbles is a cat at t, Tibbles is in P at t, there is only one cat in P and t and Tib ≠ Tibbles' is true, but so also is the conjunction 'Tib is a cat (simpliciter), Tib is in P at t, Tibbles is a cat (simpliciter), Tibbles is in P at t, there is only one cat in P at t and Tib ≠ Tibbles'.

I was careful not to distinguish these two conjunctions previously. But it is in fact only if the latter conjunction is acceptable that one needs to appeal to the distinction between restricted and unrestricted quantification to deal with the argument. For if only the former conjunction is acceptable (which must be the case if the concept of a cat is *maximal*, in the sense that no proper part of a cat is a cat) the argument can be met by interpreting 'There is only one cat in P at t' as meaning 'For some x, x is a cat (simpliciter) and x is in P at t and for every y, if y is a cat (simpliciter) and y is in P at t then x is the same cat as y'.

However, I do not think that this shows that one can in fact deal with the Tib/Tibbles example without accepting the distinction between restricted and unrestricted quantification, because I see no reason to suppose that the concept of a cat *is* maximal (here I agree with an implicit assumption of Geach's 'Paradox of the 1,001 cats' version of the Tib/Tibbles argument). But whether or not this is so, one still needs this distinction to answer the arguments for relative identity in *O&I*.

To summarise my present position, I still accept that there exist pairs of sortal terms 'F' and 'G' such that situations are conceivable satisfying the following forms of description:

(i) x is at time t the same F as y is at time t′, but x at time t is a different G from y at time t′,

(ii) x is an F at time t, y is an F at time t, x is in place P at time t, y is in place P at time t, there is only one F in P at t and x ≠ y,

(iii) x is an F (simpliciter), y is an F (simpliciter), x is in place P at t, y is in place P at t, there is only one F in place P at t and x ≠ y.

But I no longer think that the existence of such situations establishes that identity under a sortal concept is relative (even in *O&I* I did not think that the mere existence of situations of type (i) established this). I do, however, endorse the pair of distinctions (a) and (b) above, which one must also endorse if one accepts the relative identity thesis, and I believe that one is obliged to do so if one accepts the conceivability of situations of type (iii) (though not if one merely accepts the conceivability of situations of types (i) and (ii)).

University of Birmingham,
P.O. Box 363, Birmingham B15 2TT

© HAROLD W. NOONAN 1986

FACTIVES, BLINDSPOTS AND SOME PARADOXES

By BERNARD LINSKY

IN 'Fitch's Factives' (ANALYSIS 44.4, October 1984, pp. 153-8) J. J. MacIntosh generalizes an argument originating with F. Fitch which has recently reappeared in a new guise. I wish to show that MacIntosh's version is not general enough, and that when carefully viewed this argument is seen to be at the heart of several familiar epistemic paradoxes; Moore's Paradox, the Paradox of the Preface and the Propositional Liar Paradox.

MacIntosh proves that given three principles governing an arbitrary sentential operator 'D' one can derive a 'modal collapse'; that p → Dp is a theorem. The principles are:

Distrib: D(p & q) → Dp & Dq
Fact: Dp → p
Poss: p → MDp

Earlier discussions focussed on the interpretation of '... knows that ...' as such an operator D. *Poss* then says that whatever is true *can* be known. MacIntosh points out a wide range of other factive operators and argues that for none of them is *Poss* at all plausible. My generalization of the argument is in a different direction. By returning to an epistemic interpretation we see that principles weaker than *Fact* will produce the same result for such notoriously non-factive operators as '... believes that ...'.

The Fitch argument, as presented by MacIntosh, is essentially the following:

Argument I

(1)	$D(p \& {\sim}Dp) \rightarrow (Dp \& D{\sim}Dp)$	*Distrib*
(2)	$D{\sim}Dp \rightarrow {\sim}Dp$	*Fact*
(3)	${\sim}D(p \& {\sim}Dp)$	1, 2 PL
(4)	$L{\sim}D(p \& {\sim}Dp)$	3, Necessitation
(5)	${\sim}MD(p \& {\sim}Dp)$	4, Def. M
(6)	$(p \& {\sim}Dp) \rightarrow MD(p \& {\sim}Dp)$	*Poss*
(7)	$p \rightarrow Dp$	5, 6, PL

My generalization of the argument is trivial. Step (2) does follow from *Fact* but in fact is itself a weaker principle:

Ref_1: $D{\sim}Dp \rightarrow {\sim}Dp$

Such restrictions of *Fact* to particular classes of propositions may hold for operators which aren't fully 'factive'. (I call them *ref*lection principles by analogy with the notion from proof theory.) Ref_1 is a plausible principle for the clearly non-factive notion of belief. As a principle of epistemic logic $B{\sim}Bp \rightarrow {\sim}Bp$ is contentious, it requires some sort of consistency of a believer, but it is at least defensible for idealized situations where all the relevant propositions have been explicitly considered. With the substitution of 'B' for 'D', however, steps (1) to (3) of the argument become a proof of Moore's paradox. In *Knowledge and Belief [2]*, J. Hintikka interprets what G. E. Moore saw as a non-logical contradiction in saying 'p, but I don't believe it' as rather a pragmatic paradox. While the sentence $p \& {\sim}Bp$ is consistent (it *can* be true), it cannot be 'defensibly' believed, that is ${\sim}B(p \& {\sim}Bp)$ is a theorem of epistemic logic. Argument I in fact shows that this theorem only requires that belief satisfy *Distrib* and Ref_1.

Hintikka gives a very different sort of proof, one using semantic notions such as his model sets, but which can be reformulated as:

Argument II

(1)	$B(p \& {\sim}Bp) \rightarrow (Bp \& B{\sim}Bp)$	*Distrib*
(2)	$Bp \rightarrow BBp$	*BB*
(3)	$BBp \& B{\sim}Bp \rightarrow B(Bp \& {\sim}Bp)$	*Conj*
(4)	${\sim}B(Bp \& {\sim}Bp)$	*Cons*
(5)	${\sim}B(p \& {\sim}Bp)$	1–4, PL

This proof uses two further principles, both requiring logical acumen of believers:

Conj: $Bp \& Bq \rightarrow B(p \& q)$
Cons: ${\sim}B(p \& {\sim}p)$

as well as a near relative of reflection principles (just the converse of Ref_2 below):

BB: $Bp \rightarrow BBp$

Roy Sorensen [5] has introduced the term 'epistemic blindspot' for propositions that may be true but either not known or not believed. Blindspots seem to play an important role in several epistemic paradoxes, only one of which I will discuss below. One of the others is the famous 'Surprise Exam' paradox. Sorensen argues, following Binkley [1], that the paradox relies on its being possible for there to be a surprise exam, but not possible to know (or perhaps even to believe) that there is. To say that there will be an exam on a certain day will be a sentence p, to say that it will be a surprise is $p \,\&\, {\sim}Kp$. The 'paradox' arises from the obvious fact that there can be such an exam, but that one can't expect (know or perhaps believe) that it will occur as a surprise.

That the principle *Poss* is refuted by the possibility of blindspots is no accident. MacIntosh considers weakening *Poss* to rule out such counterexamples by proposing:

W(eak) Poss: $(p \rightarrow MDp) \vee (p \rightarrow \exists q(q \,\&\, {\sim}Dq))$

This principle is tautologously equivalent to:

$(p \rightarrow MDp) \vee \exists q(q \,\&\, {\sim}Dq)$

It says, quite simply, that *Poss* holds unless there are blindspots. As MacIntosh rightly points out, *W Poss* is already provable adding only the principle which states that necessity is a factive:

T: $Lp \rightarrow p$

His proof is correct but it does not show the real reason for the result. In fact one can simply read *Poss* as amounting to a denial of the existence of certain blindspots, as the following two arguments show. III Shows that if *Poss* has a counterexample then there is a blindspot of the form $(q \,\&\, {\sim}Dq)$. The converse is shown by IV:

Argument III

1	(1)	$\exists q{\sim}(q \rightarrow MDq)$	A
1	(2)	${\sim}(p \rightarrow MDp)$	1, E.I.
1	(3)	$p \,\&\, {\sim}MDp$	2, PL
1	(4)	$p \,\&\, L{\sim}DP$	3, Def M, PL
	(5)	$L{\sim}Dp \rightarrow {\sim}Dp$	T
1	(6)	$p \,\&\, {\sim}Dp$	4, 5, PL
1	(7)	$\exists q(q \,\&\, {\sim}Dq)$	6, E.G.

Argument IV

1	(1)	$\exists q(q \,\&\, {\sim}Dq)$	A
1	(2)	$p \,\&\, {\sim}Dp$	1, E.I.

(3) $\sim MD(p \,\&\, \sim Dp)$ (Argument I)
1 (4) $\sim[(p \,\&\, \sim Dp) \rightarrow MD(p \,\&\, \sim Dp)]$ 2, 3, PL
1 (5) $\exists q \sim (q \rightarrow MDq)$ 4, E.G.

So Fitch's argument proves one version of Moore's 'paradox', is useful in formulating the Surprise Examination paradox, and gives the last word on *Poss*.

A second reflection principle allows a proof of the existence of another sort of blindspot which shows up in the paradox of the Preface.

Argument V

(1) $D(\sim p \,\&\, Dp) \rightarrow D\sim p \,\&\, DDp$ *Distrib*
(2) $DDp \rightarrow Dp$ *Ref$_2$*
(3) $D\sim p \,\&\, Dp \rightarrow D(p \,\&\, \sim p)$ *Conj*
(4) $\sim D(p \,\&\, \sim p)$ *Cons*
(5) $\sim D(\sim p \,\&\, Dp)$ 1-4, PL

This second reflection principle:

Ref$_2$: $DDp \rightarrow Dp$

is another restriction of *Fact* and like *Ref$_1$* it is plausible for non-factive notions like belief. Replacing 'D' by 'B' argument V becomes a proof of the epistemic theorem $\sim B(\sim p \,\&\, Bp)$. (J. N. Williams [6] argues that Moore presents 'p but I believe that not p' as a second paradox.) It is surely possibly true that one believes p incorrectly, but not possible to believe that it is.

The belief that one has some false belief leads to a puzzle and a paradox. First, it seems that this well justified humility involves one in (doxastic) inconsistency. Suppose one believes $p_1, \ldots, p_n$. How is the belief that one has a false belief to be formulated? Not as $B(\sim p_1 \vee \sim p_2 \vee \ldots \vee \sim p_n)$ which is not jointly defensible (is 'doxastically inconsistent') with $Bp_1, Bp_2, \ldots, Bp_n$. Better as $B\exists p(\sim p \,\&\, Bp)$ which can be made true together with Bp_1, $Bp_2, \ldots, Bp_n$ in an extension of Hintikka's semantics which allows propositional variables. (In each alternative model set i, have some new proposition p_i both false and believed, where p_i is not among $p_1, \ldots, p_n$). The formula $B\exists p(\sim p \,\&\, Bp)$ should not be confused with $\exists pB(\sim p \,\&\, Bp)$ which does lead to inconsistency by a use of existential instantiation and argument V. This latter claim, however, that there is a particular false proposition that one believes oneself to believe, can be deduced from the former when one 'has a belief about which' proposition is in question. ('Knows which' is Hintikka's expression for the condition justifying the inference from $K\exists x\phi$ to $\exists xK\phi$ [3]. I use the related term for belief contexts.) Imagine a conscientious sceptic who, believing he has a false belief, re-examines his beliefs one by one, accepting each again, for his evidence has not changed, until he reaches the last candidate. He knows which one the false belief must be, he can infer $\exists pB({\sim}p \mathbin{\&} Bp)$ from $B\exists p({\sim}p \mathbin{\&} Bp)$ and finds himself in a doxastic inconsistency. How can a humble logically scrupulous believer end up in inconsistency? That is the puzzle.

The belief that one has a false belief leads to the paradox of the Preface by a step of existential *generalization* (as opposed to the instantiation considered above.) This argument has been studied by Arthur Prior and recently extended by MacIntoch [4]. Let 'Fp' abbreviate '${\sim}p \mathbin{\&} Bp$' (one *f*alsely believes that p):

Argument VI

1	(1)	$B\exists pFp$	A
2	(2)	${\sim}\exists pFp$	A
1, 2	(3)	$B\exists pFp \mathbin{\&} {\sim}\exists pFp$	1, 2
1, 2	(4)	$\exists q(Bq \mathbin{\&} {\sim}q)$	3, E.G.
1, 2	(5)	$\exists qFq$	Def F
1	(6)	$\exists pFp$	1–5 Reductio
	(7)	$B\exists pFp \rightarrow \exists pFp$	1–6 PL

It would seem that we have merely proved another reflection principle, that believing that one has a false belief makes it so. But we derive a 'paradox' by proving that the false belief must be *other* than that original belief: Since $\exists pFp$ is equivalent to

$$\exists p(Fp \mathbin{\&} p = \exists pFp) \vee \exists p(Fp \mathbin{\&} p \neq \exists pFp)$$

we derive the paradoxical result:

1 (8) $\exists p(Fp \mathbin{\&} p \neq \exists pFp)$

Where is the error in the 'paradox of the Preface'? I would favour a type-theoretic answer. The false belief shown to exist at step (6) is a second level belief, a belief that one has some false first level belief. (Hence my use of a new variable 'q' which will be of a different type from 'p'.) Strictly speaking, then, step (5) does not contradict assumption (2), there is no *reductio ad absurdum.* Of course believing that one has a false first order belief guarantees that one at least has either a false first order belief or a false second order belief, but that does not yield the paradoxical claim (8) that the error is *not* in the 'preface'. That conclusion may be reinstated by the self-referential element in the move to step (8). The purpose of this note, however, is not to finally resolve long standing paradoxes. Rather, it is to show the recurring role of blindspots and fixed points in those paradoxes.

I have made two points; one precise, the other vague. The precise thesis is that MacIntosh's argument can be generalized by noting that it does not really require *Fact* but rather only the weaker principle Ref_1. My vague point is that two 'blindspots', $(p \mathbin{\&} {\sim}Bp)$ and $({\sim}p \mathbin{\&} Bp)$, seem to appear in the formalizations of several familar epistemic paradoxes. But what is the significance of that fact?[1]

University of Alberta,
Edmonton, Alberta,
Canada T6G 2E5

© BERNARD LINSKY 1986

REFERENCES

- [1] R. Binkley, The Surprise Examination in Modal Logic', *Journal of Philosophy*, Vol. LXV, No. 5, 1968, pp. 127-36.
- [2] J. Hintikka, *Knowledge and Belief*, (Ithaca, Cornell University Press, 1962) p. 64.
- [3] J. Hintikka, 'Semantics for Propositional Attitudes', in *Models for Modalities*, (Dordrecht, Reidel, 1969) pp. 96-8.
- [4] J. J. MacIntosh, 'An Extension of a Proof of Prior's', ANALYSIS 40.2, March 1980, pp. 86-9.
- [5] R. A. Sorensen, 'Conditional Blindspots and the Knowledge Squeeze: A Solution to the Prediction Paradox', *Australasian Journal of Philosophy*, Vol. 62, No. 5, June 1984, pp. 126-35.
- [6] J. N. Williams, 'Moore's Paradox: One or Two?' ANALYSIS 39.3, June 1979, pp. 141-2.

[1] Earlier versions of this paper were read to the Banff Conference, 20 April 1985 and the Society for Exact Philosophy at Toronto, 18 May 1985. I would like to thank Ronald Clark, Hugues Leblanc and Roy Sorensen for corrections to those versions.

ON A TEMPORAL SLIPPERY SLOPE PARADOX

By D. E. OVER

THERE has been increasing interest in a range of problems which Kit Fine has felicitiously called (in unpublished work) modal slippery slope paradoxes. A good, succinct example of this type has recently been described by Nathan Salmon in his paper, 'Impossible Worlds' (ANALYSIS 44.3, June 1984, pp. 114-17).[1] My purpose here is to point out that this example, like some of the others, is paralleled by what might be called a temporal slippery slope paradox, as can be seen by considering the following principle (which parallels Salmon's principle (2′)):

(T) If a given artefact x is assembled from a set of components y, then x could be completely disassembled and later

---

[1] See also Salmon's book, *Reference and Essence* (Blackwell and Princeton U.P., 1982), Appendix I, pp. 219-59, and Graeme Forbes, 'Two Solutions to Chisholm's Paradox', *Philosophical Studies* 46, 1984, pp. 171-87.

> re-assembled from another set of components z which sufficiently overlaps y; but x could not later be re-assembled from any set of components z' which does not sufficiently overlap y.

This principle is as intuitively appealing as many of those underlying modal slippery slope paradoxes, and to see in detail how we might implicitly appeal to it, suppose that we take our watch x to a watch-maker who completely disassembles it. He then gets its components muddled up with the components of many other watches, and at a later time t' he puts a watch x' together which does not contain many of the components of our watch. We naturally reject x' as not our watch (implicitly relying on the second part of (T)), and force him to disassemble x' and to search for more of the components of our watch. Eventually, he finds nearly all of these components and reassembles at a time t_1 a watch x_1 which we acknowledge to be ours (implicitly relying on the first half of (T)). But we now seem to be on a temporal slippery slope. Perhaps our watch will have to go back to the watch-maker many times for disassembly and reassembly with a small change in component parts. In the end we might reach a time t_n when we get a watch x_n made of exactly the same components as x' put together in the same way, but which should be our watch x if the first part of (T) is applied throughout this process.

Notice that the above case is not identical with that of the famous ship of Theseus, which is continuously modified over a long period without being completely dismantled. In the latter case the slow modification of a ship may or may not bring about the existence of another ship, but in any event there is always some ship in the dry dock, and not merely a set of perhaps highly scattered components of a ship.[2] In the former case there are always times when a watch is completely disassembled and its parts perhaps used for various other purposes before most of them are collected together again and made into what may or may not be a new watch. Here there is sometimes a good reason for identifying a watch with its immediate successor in the series, sometimes a good reason for identifying a watch with some later successor in the series (as when we identified x and x_1), and sometimes there seems to be a conflict involving (T). Do we say on the basis of (T) that x_n is identical with x' or x_{n-1}, assuming that x_{n-1} is identical with x? It is true that t_{n-1} is closer in time to t_n than t' is, but this difference can be minimized by making the interval between t' and t_{n-1} very short and that between t_{n-1} and t_n very long (with the components in the relevant set used for many scattered purposes in this interval). Are we to claim that x' could *never* be reassembled after x has been reassembled as x_{n-1} at t_{n-1}? We may be wrong about the exact formulation of (T), but we do seem to presuppose that some (T)-like principle is true. And yet it is unclear how we can modify (T) so that it gives consistent answers, in line with our intuitions, to all the questions which could be asked about examples of this type.

---

[2] See E. J. Lowe, 'On the Identity of Artifacts', *The Journal of Philosophy* LXXX, 1983, pp. 220-32, for a discussion of the importance of the concept of a continuous artefact history in connection with the problem of the ship of Theseus.

An obviously central concept in all types of slippery slope paradox is that of vagueness, and it seems hard to deny that vagueness has a similar role in the modal and temporal cases (for all the important differences between the modal and the temporal). Indeed we should be sceptical of any analysis of this concept in the modal case which does not apply to the temporal one. Salmon, for example, has tried (in the work cited in footnote 1) to dissolve the modal paradoxes by holding that the accessibility relation between possible worlds is vague and not transitive, but this idea does not generalize to the temporal paradoxes and is something of a technical *deus ex machina.* One proposal that does apply to both types of paradox is that difficulties arise in all these cases through the use of *vague* or *imprecise designators*, i.e. singular terms, like 'our watch', used by speakers who have not decided whether the referents of the terms are composed of certain sets of components in some possible worlds or at some points of time.[3] Consider a use of a term like 'our watch' when its referent was *originally* assembled. At that time we would not have been definite about whether this term referred to a watch that would be assembled again at t_n after the series described above. How could we have known that this possibility would arise? We were precise enough about the referent of 'our watch' to know that it was not identical with the referent of a term used to refer to the new watch assembled at t'. But the latter term would also be imprecise to some extent – we would not settle at t' whether it referred to a watch that is reassembled at t_n. It would thus be unclear which one (if any) of two such relatively vague designators referred to a watch assembled at t_n in the described circumstances. Analogously in a modal case, one would begin with the world in which the referent of 'our watch' was *actually* constructed, i.e. the actual world. One would go on to note that being relatively definite about the assembly of the referent in this world would not necessarily mean being as definite about its assembly in some other possible worlds, and then similar points to those made in the temporal case would naturally follow.

The amount of vagueness in a singular term like 'our watch' normally causes no more trouble than the amount of vagueness in the general term 'watch'. But the existence of vague designators does explain how slippery slope paradoxes *in general* can arise, and the concept of a *regimentation* can be used to take a step towards escaping from them when they do arise (regimentations are introduced by Thomason in his important paper cited in footnote 3). We would regiment 'our watch' (to some extent) by deciding whether or not its referent was reassembled at t_n, and consistently pair this with decisions about any other relevant designators. For example, 'our watch' could be used consistently if the option chosen was a regimentation in which its referent got reassembled at t_n, while the referent of 'the watch first assembled at t''', did not. But of course possible decisions of this type cannot be the end of the story, as many questions immediately come to the fore. To what extent is it arbitrary which regimentation is chosen in these cases? Is it ever possible for terms referring to separate artefacts to have the same regimentation at some points of time or in some (non-actual) possible worlds (so that, for instance, our watch and the distinct watch first assembled at t' could *both* be thought of, under some one regimentation, as reassembled at t_n)?

---

[3] The concept of a *precise designator* was introduced, on perfectly natural grounds independent of the paradoxes, by Richmond H. Thomason in 'Identity and Vagueness', *Philosophical Studies* 42, 1982, pp. 329–32.

How should (T) and other temporal and modal principles be modified given the existence of vague designators? Would our intuitions about these principles have to be changed in the light of the modifications? It must be admitted that these are difficult questions, but then the depth and generality of slippery slope paradoxes must be fully appreciated before a complete solution is possible.[4]

Sunderland Polytechnic,
Sunderland SR1 3SD

© D. E. OVER 1986

---

[4] I should like to thank Graeme Forbes and Nathan Salmon for discussion and correspondence. Very special thanks are due to Kit Fine for helping me to see the real seriousness of modal slippery slope paradoxes.

ON A PARADOXICAL TRAIN OF THOUGHT

By DORIS OLIN

IN a recent paper ('A Paradoxical Train of Thought', ANALYSIS 44.4, October 1984), Martin Hollis presents a seemingly paradoxical line of reasoning, the core of which may be put thus. A and B each think of a number and whisper it privately to C. C then announces 'You have each thought of a different positive number. Neither of you can determine whose is greater.' A, who has picked 157, reasons as follows. 'B cannot have picked 1, since, if he did, he would be able to infer that I had picked a greater number; for the same reason, he knows that I did not choose 1. But, given this, it also appears that no one chose 2. For if B chose it, he would know that my number is greater; and if I chose it, I would know that his is greater. Similarly for 3. And so on. Hence, the announcement cannot be satisfied.'

Hollis's paradoxical train of thought, I suggest, follows a familiar route. It is a version of the well known prediction paradox, which is typically expressed in terms of a surprise examination. The basic structure of the prediction paradox is as follows. An audience is informed that a certain state of affairs obtains (will obtain), and also that they are (will be) ignorant of the truth concerning some particular aspect of that state of affairs. (For example, in one familiar version, the audience is informed that an examination will be given during a certain week, and that they will not know beforehand on exactly which day it will be given.) The paradox consists in the fact that one appears to be able to prove, by impeccable reasoning, that the announcement cannot be satisfied while, in fact, it clearly can be.

There are, it should be noted, some interesting differences between Hollis's paradox and the standard versions of the prediction paradox. For one thing, the standard versions involve a finite number of possibilities (for example, the five days of the week), while in Hollis's version there are infinitely many possibilities to be eliminated (B chose 1, B chose 2, etc.). Nevertheless, there is nothing *substantially* new here. The proof of this is that Hollis's version falls to the same solution which resolves the familiar versions of the paradox. (This solution is presented in my paper 'The Prediction Paradox Resolved', *Philosophical Studies*, 44, September 1983.)

Let us see, briefly, how the solution applies. The paradox generating argument fails, I suggest, at the very first step:

> If B had picked 1, he would be able to infer that A had picked a greater number.

Suppose that B had picked 1. Would he be justified in believing that A had chosen a greater number? Well C, normally a reliable fellow, said:

> (P) A and B each picked a different positive number.
> (Q) If A's number is larger than B's, neither of them is entitled to believe this; similarly if B's number is larger than A's.

In order to infer that A's number is greater, B would have to rely on part (P) of the announcement. But would B in fact be justified in believing (P) if he had chosen 1? I think not. Note, first, that he would be justified in believing (P) only if he were also justified in believing (Q), for his evidence for one is no better than his evidence for the other. However, if B were justified in believing both (P) and (Q), then, since he chose 1, he would also be justified in believing:

> A's number is greater than mine and I am not now justified in believing that A's number is greater than mine.

But this is impossible. One can never be warranted in believing a proposition of the form 'p and I am not now justified in believing that p'. The upshot is that if B had chosen 1, he would not be entitled to believe both (P) and (Q); further, since he has no reason to prefer either part of the announcement, he would not be warranted in believing just (P). The paradox generating argument thus collapses at the very first step.

The solution, as thus far presented, depends, of course, on the fact that there is equally good evidence for (P) and (Q). If we revise the situation so that there is better evidence for (P), the solution will be a bit more complicated. Briefly: if either party were to choose 1, he would still be warranted in believing (P), since there is better evidence for (P); and he could therefore infer that the other person has chosen a greater number. However, the argument can now be stopped at the next stage:

> If B had chosen 2, he would be able to infer that A had chosen a greater number.

The reasoning underlying this step is that B would be able to infer that A had not chosen either 1 or 2; and this inference would be based, presumably, on B's justified belief in both (P) and (Q). However, in this revised situation, if B had chosen 2, he would not be justified in believing both (P) and (Q); for if he were, he would also be justified in believing:

> A's number is greater than mine and I am not now justified in believing that A's number is greater than mine.

But this is impossible.

Thus, Hollis's paradoxical train of thought is brought to a halt.

Glendon College, York University,
2275 Bayview Avenue,
Toronto, Ontario, Canada M4N 3M6

© Doris Olin 1986

A PARADOX DERAILED: REPLY TO HOLLIS

By Michael Kinghan

In Analysis 44.4, October 1984 Martin Hollis invites solutions from the readership of this journal to 'A Paradoxical Train of Thought'. I review his paradox and solve it below.

Persons A and B each choose a positive integer and communicate their choices privately to C. C then informs them, truly, that they have chosen different integers, and he predicts that neither of them can work out who has chosen the larger. We are told that A has chosen 157.

It seems to A that a paradox ensues:

> I picked 157 and have no idea what B picked. So, assuming that he indeed chose a different positive whole number, C is right. I wonder what follows from that. Well, clearly B did not choose 1, as he would then be able to work out that mine is the greater; and by the same token he knows that I did not choose 1. So he did not choose 2, since he could then use the previous reasoning to prove that my number is greater. Similarly he can know that I did not choose 2 either. With 2 out of the way, I can infer that he did not choose 3; and he can infer that I did not choose 3.

By "mathematical induction" it is apparent to A that the continuation of this reasoning eventually eliminates every positive integer as being either B's choice or his own. So if C's prediction is true, then neither of them can have chosen any integer, which is absurd. But so, Hollis urges, is the contrary supposition:

> 'Suppose', A ponders, 'C was wrong to say that neither of us can work out who has chosen the greater number . . . If I can, it can only be by using the mathematical induction to place B with the greater. But in that case, he can also use it to place me with the greater. So if C is wrong, we can each prove the other to have the greater. That is absurd too. So C is right after all'.

Suspicion is naturally drawn to the "mathematical induction", but ere we trouble to appraise that argument we can at once reproach A with a spectacular combination of non-sequiturs.

Initially he affirmed that C's prediction was correct, given that the integers are indeed different. He affirmed this on the basis that he has chosen 157 and had no idea what B had picked, the implicit point being that a positive integer known only to be different from 157 may equally, for all one can possibly tell, differ upon the small side as upon the larger. But of course, *A's* inability to solve the problem does not entail that C is right, for C has predicted that *neither A nor B* can solve it. Having no idea what B has chosen, A is unable to solve the problem and, for all he knows, B has chosen 1: in which case, *B* could solve the problem, just as A next observes.

Having been converted to the absurdity of his initial judgement by the "mathematical induction", A is then troubled to discover that the contrary supposition fares no better: for if, contrary to the prediction, he can work out whose choice is the greater it can only be by using the "mathematical induction" to place B with the greater; and then B could also use it to the opposite result.

Here A has forgotten that the "mathematical induction" supposedly enables him to eliminate any choice by B smaller than his

P4708

own only on the hypothesis that C's prediction is true. But he is now supposing that prediction is false. In that case he cannot apply the "mathematical induction" to place B with the greater number, so there is no purported reason why B cannot have chosen 1 and discerned, falsifying the prediction, that the greater number must be A's.

Further, A by now considers that a case in which he can apply the "mathematical induction" is one in which he can place B with the greater number, whereas the only case in which he can purportedly apply it was supposed to be one under which it would follow that nobody chose any integer. But forgetfulness that the "mathematical induction" has the truth of the prediction as its major hypothesis obscures from A the fact that, whatever conclusions were thereby inferred by whatever parties, their arguments could not either severally or jointly afford a *reductio ad absurdum* of the prediction's falsity: it is tautologous that if the prediction is false then, if it is true, then A chose the larger integer and so did B — and neither of them made any choice.

Finally, A has transposed his first non-sequitur. From the supposition that C is wrong, i.e. that *either A or B* can solve the problem, it does not in any case follow that *A* can solve it, *a fortiori* does not follow that there is a way in which A can solve it which could also be applied by B to the opposite result. Once again the prediction would be false in the event that B has chosen 1 and uses the obvious non-inductive reasoning which would in that case be available to him, but not to A, to place A with the larger number.

In sum, A had no plausible basis for initially upholding C's prediction and, having repudiated that view on the basis of the "mathematical induction", had finally no plausible basis for reversion.

Can any semblance of a paradox be salvaged? Yes. The "mathematical induction" is purportedly a *reductio ad absurdum* of C's prediction and we have yet to challenge that argument. The character A made no assumptions about B's choice and would have been an idle vehicle for the entertainment of assumptions outrunning his supposed knowledge: it was evidently Hollis' aspiration to devise a paradox inherent in the information of that character. However, in the execution of his first non-sequitur, A was not to be faulted in his apparent appreciation that, when a positive integer is known only to be other than 157, there is no knowing whether it is smaller than that or larger, and that therefore he could not solve the problem. The same would hold, of course, had his chosen integer been anything other than 1. So if we relinquish the aspiration imputed to Hollis and assign to *both* parties choices other than 1, the appearance of paradox revives; for on such an assignment neither party can solve the problem and the prediction is true, in spite of the "mathematical induction". One's only business, in this case, is to expose the error of that argument.

Here A accomplished his subtlest fallacy, the fallacy of epistemic distribution: if N knows that p implies q, then p implies that N knows that q.

Let us add to Hollis' story the assumption that B's choice is other than 1: so the prediction is true. I find too that both brevity and clarity are assisted by the fancy that B is none other than myself. Let P be C's prediction. A wonders what follows from P:

> 'Well, clearly Kinghan did not choose 1, as he would then be able to work out that mine is the greater; and by the same token, he knows that I did not choose 1.'

Here, A takes himself to have established that, if P is true, then I know that he did not choose 1: this on the assumption, which we may grant, that my wits are at least as good as his. Accordingly, he should already have established that, if P is true, then *he* knows that *I* did not choose 1. But of course he has not done that, else he need only reflect that he has no idea whether I have chosen 1 to conclude forthwith that P is false and that one of us can determine whose choice is the greater, and plainly his ignorance on that point has no such implication. He has already established only that, if P is true, then I did not choose 1. So he knows that much, and might fairly regard as established, 'by the same token', *my* knowledge that, if P is true, then *he* did not choose 1. But as we have remarked, he takes himself to have established that, if P is true, then I know that he did not choose 1.

His confusion embodies the distributive fallacy. The sum of his entitlement, by sameness of token, is our common knowledge that, if P is true, then the other did not choose 1. From that he might further be allowed to infer that, if either of us *knew* that P was true, then that one would also know that the other had not chosen 1. But our present assumptions exemplify the evident possibility that P should be true while neither of us knows it: neither of us has chosen 1, so P is true and, for all that either of us can tell, the *other* has chosen 1 and thereby determined that his choice must be the smaller.

Now to ground his induction, A must proceed to show that, if P is true, then I cannot have chosen an integer all of whose predecessors are already eliminated on the supposition of P. Only 1 has been eliminated on that supposition. So to eliminate my choice of 2, A takes my knowledge that he has not chosen 1 for a previously established consequence of P. Given that, then if P is true I cannot have chosen 2; for if I had I would be able to see that A's choice must be greater than mine. But as we have noted, P does not have that consequence. And as we have further noted, in order to know that A had not chosen 1 I should need to know that P was true. But had I chosen 2 I could not know that P was true, because, as again was noted, if either of us has not chosen 1 then he cannot know that the other has not done so and thus placed himself with the smaller number.

A does not take inductive flight immediately upon disposing of my choice of 2, but elliptically invokes first principles at 3, 4 and 5. It will be seen that each such invocation must be reckoned to involve an additional perpetration of the distributive fallacy.

42 Argyle Street,
Oxford OX4 1FS

© Michael Kinghan 1986

ON THE LIMITS OF SCIENCE

By George N. Schlesinger

NICHOLAS Rescher in his latest book declares that 'perfected science is a mirage; completed knowledge a chimera'. He advances a formal proof to 'provide some suggestive stage-setting for the more concrete rationale of the imperfectibility of science'. He begins by formulating four elementary theorems of epistemic logic many would readily accept:

(K1) It is impossible to know anything that is false. Thus if a proposition is known then it is true:

$$Kp \supset p$$

(K2) A conjunction cannot be known unless each of its conjuncts is known. Thus, if a conjunction is known then both of its conjuncts are also known:

$$K(p \,\&\, q) \supset (Kp \,\&\, Kq)$$

(K3) Nobody would claim that already now we know everything there is to be known; there is a vast number of true propositions that at present are unknown to us:

$$(\exists p)(p \,\&\, {\sim}Kp)$$

(K4) All truths are knowable:

$$p \supset PKp$$

Rescher then proceeds to show that these four theses are inconsistent:

1. $K{\sim}Kp \supset {\sim}Kp$ substituting Kp/p in (K1)
2. ${\sim}(K{\sim}Kp \,\&\, Kp)$ from 1 by the definition of '$\supset$'

3.	$K(p \,\&\, \sim Kp) \supset (Kp \,\&\, K{\sim}Kp)$	substituting Kp/q in (K2)
4.	$\sim K(p \,\&\, \sim Kp)$	from 2, 3 by Modus Tollens
5.	$N{\sim}K(p \,\&\, \sim Kp)$	from 4 by the rule of necessitation of valid sentences
6.	$\sim PK(p \,\&\, \sim Kp)$	from 5
7.	$\sim(p \,\&\, \sim Kp)$	from 6, (K4) by Modus Tollens
8.	$(p){\sim}(p \,\&\, \sim Kp)$	from 7 by generalization
9.	$\sim(\exists p)(p \,\&\, \sim Kp)$	from 8

Thus (K1)-(K4), which lead to a contradiction (as 9 contradicts (K3)), are incompatible. Hence Rescher concludes:

> This argumentation shows that in the presence of the relatively unproblematic principles (K1)-(K2), the thesis that all truths are knowable (K4) entails that all truths are known, that is not – (K3). Since the latter thesis is clearly unacceptable, the former must be rejected. We must concede that some truths are unknowable: $(\exists p)(p \,\&\, \sim PKp)$. (*The Limits of Science*, University of California Press, 1984, p. 150)

Rescher adds that Richard Routley has produced what is 'a close cousin of this argumentation', also yielding the same important result.

Unfortunately it is not granted to us to fathom with the aid of elementary epistemic logic what knowledge will or will not elude us till the end of days. Closer inspection shows Rescher's argument to be fallacious. To see this, we should ask: what precisely does his symbol 'K' stand for? If we are to judge by theorem (K3) then unquestionably 'K' signifies 'It is known *at present* that ...' since what that theorem states is something everyone will concede to be obviously true, namely, that at the present moment none of us knows everything. (K3) is surely not intended to assert that *at no time* is everything known, since that is by no means evidently true (and Rescher himself admits it to require a seven-step proof). But if so, the final conclusion '$(\exists p)(p \,\&\, \sim PKp)$' states nothing more noteworthy than that there is some true proposition, which *at the present time*, is in principle unknowable.

To put it slightly differently, let

$K{*}p$ = The probability that (at some time in the future) *p* is known, is *not* zero.

Clearly, then, the principle of the imperfectability of science is to be symbolized as

$$(K{*}3)\,(\exists p)(p \,\&\, \sim K{*}p)$$

meaning that there is at least one proposition of which it is absolutely certain that it will *never* be known (i.e. the probability *is* zero that will ever be known). Those, however, who do believe in the perfectability of science subscribe to

$(K*4)\ p \supset K*p$

which is the denial of (K*3).

Now step 4 of the derivation states 'It is false that it is known *now*, both that *p* and that it is false that *p* is known now'. Consequently, what Step 6 states is that it is necessarily false (or logically impossible) that K(...), i.e. that it be known *now* that (...). Bearing this in mind, we can see it clearly that Step 6 and (K*4) cannot be combined to yield Step 7, by Modus Tollens, since the operator K* does not appear in Step 6.

It seems that Professor Rescher has presented us with a remarkable argument of the kind that is rarely found in philosophical literature, an argument which can be refuted through any number of easily performed experiments. I myself have achieved very clear results with the aid of a fairly simple experimental test conducted in the context of the following stipulations: the variable *p* is restricted to range over propositions describing the colour of any marble in an opaque box B; 'Kp' stands for 'It is known to my friend that *p*' (leaving the question 'when' open, like Rescher).

The actual experiment consisted in placing three marbles in B before *t*, and preventing my friend from acquiring a knowledge of the colour of at least one of the marbles before *t*; that is, I made sure that $(\exists p)(p \,\&\, {\sim}Kp)$ was true at any time before *t*. This gave me (K3). (K1)-(K2) were assumed as given. Adding (K4), I followed Rescher and arrived at the conclusion that there is at least one true description of the colour of one of the marbles which my friend will at no time know. At *t* I opened B for my friend to inspect all the marbles. It turned out that he ceased to be ignorant of the colour of any one of the marbles.[1]

University of North Carolina,
Chapel Hill, North Carolina 27514, U.S.A.

© GEORGE N. SCHLESINGER 1986

---

[1] I am grateful to the Editor for his valuable comments on an earlier version of this note.

THING STATEMENTS AND APPEARANCE STATEMENTS

By A. J. DALE

IN a recent exchange of papers in this journal Chisholm's argument in [1] to the effect that no statement referring solely to appearances is entailed by an ordinary thing statement has been challenged by M. J. Maloney [4] and defended by E. R. Kraemer [2]. I shall argue that the phenomenalist may escape from Chisholm's argument and Kraemer's defence of it by maintaining the necessity of certain bridging conditionals.

Chisholm argued that the simple thing statement (P) 'this thing is red' entails no appearance statement by itself. Only when taken in conjunction with some other thing statement referring to observation conditions will an appearance statement (R) 'redness will be sensed' be entailed. Maloney argues that since the conjunction of the simple thing statement and the thing statement referring to observation conditions is itself a thing statement (on the reasonable assumption that the conjunction of two thing statements must itself be a thing statement) there is indeed an appearance statement entailed by a thing statement.

Kraemer has challenged Maloney's argument by examining the thing statement referring to observation conditions that Chisholm uses:

> (Q) This is perceived under normal conditions: and if this is red and perceived under normal conditions, redness will be sensed ([1], p. 192).

Kraemer then gives a faulty argument to the effect that the conjunction of P and Q under a certain interpretation of 'normal conditions' together with the statement (H) 'the observer is hallucinating yellow sense-data' does not entail R. His stated reason for denying the entailment is that 'normal conditions' must exclude such hallucinatory sensings and if an interpretation of the phrase 'normal conditions' fails to exclude such sensings (P & Q) is consistent with H. But this is not a good reason for denying the entailment: that (P & Q) does entail R is a simple matter of propositional logic no matter how 'normal conditions' is interpreted. So (P & Q & H) entails R. (There are no subjunctive conditionals here!) He should rather have said that we would not accept Q unless 'normal conditions' is interpreted as excluding such hallucinatory sensings.

Kraemer's main objective, however, is that Q is a *mixed* statement containing references both to physical objects and to appearances and does not count as a pure thing statement. Thus (P & Q) is not itself a pure thing statement and so no appearance statement has been shown to follow from a pure thing statement.

What Kraemer overlooks is that if the mixed statement is itself *necessary* it may be dropped from a conjunction without affecting the entailment of any statement entailed by that conjunction. There are *consequentiae materiales* as well as *consequentiae formales*: 'John is a bachelor' entails 'John is unmarried' no less than 'John is a bachelor and if John is a bachelor then John is unmarried' entails 'John is unmarried'. The dropping of the 'mixed' statement 'if John is a bachelor then John is unmarried' fails to affect the entailment. As it stands, however, Q is obviously not a candidate for necessity status but it can be dismantled into two components:

Q_1 This is seen under normal conditions
Q_2 If this is red and seen under normal conditions then redness will be sensed.

Here Q_1 is a thing statement and Q_2 a mixed statement. Now it is certain that (P & Q_1 & Q_2) entail R: if Q_2 is necessary (P & Q_1) will entail R. So a thing statement will entail an appearance statement after all. But is Q_2 necessary? I do not pretend to know the answer but an affirmative answer is not uncongenial to phenomenalists and others explicitly accept the necessity of Q_2 (see [3], p. 61). This does not imply that the phenomenalist's problems are over since there may of course be further difficulties over the analysis of the thing sentence Q_1 with its reference to normal conditions. Chisholm after all distinguished between 'ordinary' thing statements and thing statements that refer to observation conditions though he did not stress this distinction. His contention then that no appearance statement is entailed by any *ordinary* thing statement is left untouched by the above considerations.

University of Hull,
Hull HU6 7RX

© A. J. DALE 1986

REFERENCES

[1] R. Chisholm, *Perceiving*, New York, 1957.
[2] E. R. Kraemer, 'Phenomenalism and Observation Conditions', *Analysis* 44.3, June 1984, pp. 140-43.
[3] C. McGinn, *The Subjective View*, Oxford, 1983.
[4] M. J. Maloney, 'Chisholm's Objection to Phenomenalism', *Analysis* 42.1, January 1982, pp. 25-6.

A NOTE ON DAVID LEWIS'S REALISM ABOUT POSSIBLE WORLDS

By MARGERY BEDFORD NAYLOR

DAVID Lewis has given the following argument that there are possible worlds.

> I believe that there are possible worlds other than the one we happen to inhabit. If an argument is wanted, it is this: It is uncontroversially true that things might have been otherwise than they are. I believe, and so do you, that things could have been different in countless ways. But what does this mean? Ordinary language permits the paraphrase: there are many ways things could have been besides the way that they actually

> are. On the face of it, this sentence is an existential quantification. It says that there exist many entities of a certain description, to wit, 'ways things could have been'. I believe that things could have been different in countless ways; I believe permissible paraphrases of what I believe; taking the paraphrase of its face value, I therefore believe in the existence of entities which might be called 'ways things could have been'. I prefer to call them possible worlds. (D. Lewis, *Counterfactuals*, Blackwell 1973, p. 84)

If we accept this argument, then why should we not accept the following argument that there are *impossible* worlds?

It is also uncontroversially true that some things might *not* have been otherwise than they are. The chair that is in fact blue could not have been both blue and red in the same respect and at the same time. I believe, and so do you, that things could *not* have been different in countless ways. But ordinary language permits the paraphrase: there are many ways things could *not* have been besides the way that they actually are. On the face of it, this sentence is also an existential quantification. It says that there exist many entities of a certain description, viz., 'ways things could *not* have been'. I believe permissible paraphrases of what I believe. Taking the paraphrase at face value, I therefore believe in the existence of entities which might be called 'ways things could *not* have been'. I call them *impossible* worlds.

Syracuse University,
Syracuse, New York 13210, U.S.A.

© MARGERY BEDFORD NAYLOR 1986

EHRING'S THEORY OF CAUSAL ASYMMETRY

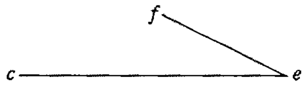
By GREGORY BASSHAM

IN a recent paper ('Causal Asymmetry', in *The Journal of Philosophy*, LXXIX, 12 (December 1982): pp. 761–74; subsequent page references are to this article) Douglas Ehring proposes an account of causal priority or asymmetry based on the 'circumstantial' character of causal relations. The purpose of his paper, he says, is to provide an 'adequate reconstruction' of causal priority which will 'apply to all empirically possible worlds' (pp. 764, 771). In the present paper I want to suggest various reasons for thinking that Ehring's account fails.

I

Much of the theoretical attraction of the traditional Humean account of causal priority lies in its clarity and simplicity: if an event *c* causes another event *e*, *c* is temporally prior to *e*, *simpliciter*. (The difficulty with this temporal-priority view, as Ehring notes, is that it cannot countenance the (physical or logical) possibility of simultaneous or backwards causation.) Alternative accounts of causal priority, by comparison, are typically cluttered and obscure; Ehring's theory is no exception to this. Still, the basic intuition underlying Ehring's account may be summed up simply enough: on a particular occasion a given event *c* causes another event *e* just in case there are certain additional conditions or events, at least some of which are causally connected to *e* but not to *c*. Thus if we assert, for example, that a particular perceptual experience is caused in part (in ordinary circumstances) by a shooting star, we imply, on Ehring's account, that there is some set of events or conditions – the fact, for instance, that it is night rather than day – that is causally related to the former but not to the latter. If we symbolize the shooting star as *c*, the perceptual experience as *e*, and the nocturnal condition as *f*, this may be diagrammed as follows:

(I)



On Ehring's (as yet unrefined) analysis, the causal priority of *c* to *e* in (I) follows straightway from the causal independence of *e*'s multiple causes.

This, as I say, is the rough intuitive picture. More precisely, Ehring suggests:

> *c* is *causally prior* to *e* if and only if *either*
>
> (A) (1) *c* and *e* are causally connected, and
> (2) there is some condition of the causal connection between *c* and *e* which is not connected causally to *c* and is causally connected to *e*, and there is no condition causally connected to *c* but not to *e*, or
>
> (B) *c* is causally connected to some event *f* and *c* is a direct condition of a causal connection between *f* and *e*, and *f* is causally prior to *e* (p. 770).

Two key notions in this account require clarification before we can consider its adequacy. The first, that of a 'causal connection', is relatively unproblematic: some event *c* is *causally connected* to another event *e* just in case *c* causes *e* or *e* causes *c*. The second notion, that of a 'condition of a causal connection', is unfortunately less tractable.

Roughly speaking, for Ehring, an event *f* is a *condition of the causal connection* (CCC, for short) of events *c* and *e* if the causal connection of *c* and *e* is counterfactually dependent upon *f*; that is, it is a CCC if, if *f* had not occurred, *c* and *e* would not have been causally connected. More precisely, however, since Ehring clearly needs to exclude, as possible CCCs, pre-empting and overdetermining causes on the one hand, and joint effects of a cause on the other:

> *f* is a direct condition[1] of the causal connection between *c* and *e*, where *c* and *e* are particular events, if and only if
>
> (1) (a) *c*, *e*, and *f* occur, and *c* and *e* are causally connected,
> (b) *f* is a member of a set of events ($f, d_1, d_2, \ldots, d_n$) each of which occurred, such that the causal connection between *c* and *e* is counterfactually dependent upon each member of the set in the absence of the other members, where this set includes all and only nonredundant[2] members, and
>
> (2) there is no event *g* upon which *f* counterfactually depends which is not a member (either redundant or not redundant) of the set ($f, d_1, d_2, \ldots, d_n$) (p. 767).

II

Does the foregoing provide an 'adequate reconstruction' of causal priority which will 'apply to all empirically possible worlds'? I have my doubts, including the following.

(1) Ehring criticizes a theory of causal priority proposed by David Sanford on the ground that it succeeds in guaranteeing asymmetry in all cases only if one is prepared to rule out, by conceptual fiat, a possible causal pattern: that in which a causal pair is not part of a larger causal chain. But here a *tu quoque* is surely in order. For nothing in our concept of causation entitles us to rule out *a priori*, as Ehring's account does, either (i) the production of an effect from a single causal event or (ii) the production of an effect entirely by causes which are themselves causally connected. These seem *ad hoc* restrictions, required not by our concept of causation but by the exigencies of a particular theory.

(2) Ehring's account presupposes that causal relations must be asymmetrical at least in the following respect: that if an event *c* causes an event *e*, then *e* does not cause *c*. But why, once the possibility (logical and perhaps physical) of backwards causation is admitted, should this presumption be retained? The following picture, in which an event *c* (in part) causes an event *e*, and *e* in

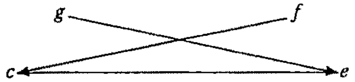
---

[1] The notion of a 'direct CCC', for present purposes, may be defined by example: thus *f* in (I) is a direct CCC. Ehring defines an 'indirect CCC' as follows: *g* is an indirect condition of a causal connection between *c* and *e* if and only if *g* is a direct condition of a causal connection between some direct condition *f*, of a causal connection between *c* and *e*, and either *c* or *e* (p. 769).

[2] 'A member of the set is redundant if the elimination of that member does not require the elimination of some other member' (p. 767).

turn (partially) backward-causes *c*, seems at least to be a conceptual possibility:

(II)



What this example shows, I think, is that asymmetry (in the sense specified) is not a necessary feature of causal relations, as Ehring assumes. 'Causal priority' and 'causal asymmetry' are not, as Ehring treats them, interchangeable terms.

(3) Finally, Ehring's account of causal priority involves a further presumption that no account of causal priority, as such, should involve: that every event in the universe is not causally connected to, and counterfactually dependent upon, a first cause. How does Ehring's account presuppose this? Suppose that there is a first cause *c*, upon which every event in (the history of) the universe is counterfactually dependent. Suppose, further, that we want to establish the causal priority of *c* to some event *e* (e.g. the ignition of the first star). It is evident that clause (A) of Ehring's definition is not applicable, since in the present example there is *no* CCC of *c* and *e* which is not connected causally to *c* and is to *e*. Nor does clause (B) apply, since Ehring excludes (in clause (2) of his definition of a CCC) '*joint effects* of a cause *c* from counting as a condition of the causal connection between *c* and some other effect *e*' (p. 768; my italics). But every event, *ex hypothesi*, is causally connected to, and counterfactually dependent upon, *c*. Thus *every* event (given Ehring's contention that 'is a cause of' is transitive (p. 765)) is a joint effect of *c*, and there is no event which can count, as clause (B) requires, as a CCC of *c* and *e* (or *c* and any other event in the universe). Thus on Ehring's analysis we would have to conclude, contrary to our hypothesis, that *c* is not causally prior to *e* or to any other event.

University of Oklahoma,
455 West Lindsey Street,
Norman, Oklahoma 73019, U.S.A.

© GREGORY BASSHAM 1986

CLOSED CAUSAL LOOPS, SINGLE CAUSES, AND ASYMMETRY

By DOUGLAS EHRING

GREGORY Bassham's comments in 'Ehring's Theory of Causal Asymmetry' provide a welcome opportunity to clarify and to revise views presented in my 'Causal Asymmetry' (Gregory Bassham, 'Ehring's theory of causal asymmetry', ANALYSIS, this issue, and Douglas Ehring, 'Causal asymmetry', *Journal of Philosophy* 79, 1982, pp. 761–74). Bassham presents what I take to be three principal problem cases in challenge to my analysis: closed causal loops in which asymmetry fails, single cause cases in which there is no plurality of causes, and a 'first cause' case in which *all* events in the universe are causally connected to a first cause. I will begin my treatment of these counterexamples with his closed causal loop example.

Bassham argues that whereas I hold that 'causes' is asymmetrical, I *ought* to allow for the possibility of cases in which an event c causes an event e and e causes c. While I would agree that closed causal loops ought not be ruled out, I do not accept Bassham's description of these cases. Rather than saying that a closed causal loop between c and e consists of c's causing e and e's causing c, I would say that c and e are 'mutually causally dependent', a symmetrical relation. This way of describing the situation preserves the asymmetrical character of 'causes' without loss of descriptive capacity. In fact, my analysis was intended to specify, implicitly, conditions under which c and e stand in a closed causal loop, viz.: c and e are mutually causally dependent just in case c and e are causally connected and c is neither causally prior to nor causally posterior to e. Nevertheless, Bassham's point does call for one revision in my characterization of 'causal connection'. In my paper it was said that c is causally connected to e if and only if c causes e or e causes c (Ehring, p. 764). Given the asymmetrical nature of 'causes' and the possibility of mutual causal dependence, a third disjunct must be added: c is causally connected to e if and only if c causes e, e causes c, or c and e are mutually causally dependent. Truth conditions for the statement 'c is causally connected to e' would establish that the disjunction is true without determining which disjunct is true. In the account of causal priority that I presented, these truth conditions were not offered, i.e. 'causal connection' was treated as a primitive notion. In that the important logical features of the causal connection relation, for the purposes of a theory of causal asymmetry, are its symmetry and non-transitivity, this modification calls for no further change in the analysis.

Bassham presents a problem case consisting of an effect e being brought about by a *single* cause c. In my paper 'Causal Asymmetry' I briefly touched on such cases with the following remark: 'In a logically possible world in which events were causally connected in a chain "one event thick" this definition (of causal priority) would fail to apply if the causal pairs in such a chain possessed causal direction' (ib. footnote 19, p. 772). I am, in fact, inclined to think that causal pairs in such a chain would lack causal direction altogether, i.e. the members of each pair would be mutually causally dependent. However, in that my intuitions are not firm on this point, I will now consider how my account can be extended to such cases. One possible way of effecting this extension is by reference to what might be considered to be an important sub-class of events – negative events or 'unchanges'. The class of causal relata could be widened to include such non-occurrences. On this approach it will turn out in general that in addition to the positive events which condition causal connections, there are negative events which play this role. For example, the causal connection between striking the match and its lighting will be conditioned not only by the presence of oxygen but the absence of moisture. Given this inclusion of negative events into the class of causal relata, one would be hard pressed to describe a case in which an effect is the result of a single cause, be it a positive event or a negative event. This way of handling the objection, however, leads to its own problematic consequences. Allowing negative events as causal relata would seem to force us to countenance an unacceptable increase in the number of causal relations in the world; e.g. every positive event which is caused will have an infinite number of causes and there will be a whole new class of causes and effects, negative events caused by negative events. Secondly, generally we ought to avoid admitting 'negative' entities into our ontology.

There is, however, a better, ontologically neutral strategy involving 'negative events', a strategy which does not commit us to the existence of negative events or to the over-proliferation of causal relations. We do not so much require the existence of negative events but the truth of certain counterfactual statements of roughly the following form: 'the causal connection between c and e counterfactually depends upon the non-occurrence of some event f'. Statements of this sort can be true even if ontological status is not granted to 'non-occurrences'. We can then read 'there exists a negative event non-f which is a condition of the causal connection (CCC) of c and e' as standing in for a counterfactual statement of roughly this form (for complications see ib. p. 767). Secondly, in order to bypass any commitment to the causal efficacy of non-occurrences, we can stipulate that events which stand in causal connections (including causes, effects, 'mutual determiners', and CCC's) are restricted to *occurrent* events (in ib. footnote 8, p. 765, I did not exclude "unchanges" from the class of causal relata). With these two points in mind, concepts analogous to 'causal connection' and 'CCC' are devisable by deleting from the truth-conditions of these latter concepts the requirement that the relata are occurrent events, and replacing it with a clause *excluding* occurrent events as possible relata. Let us call the first analogue concept 'quasi-causal connection' and the second 'quasi-CCC'. With this terminology we can say, for example, that the *absence of moisture* was a quasi-CCC between the striking of a match and its lighting on a given occasion. This way of describing the situation will *not* commit us to either the existence of this 'absence' or its causal efficacy. The analysis of 'causal priority' can then be revised as follows in order to take account of Bassham's single cause case:

c is causally prior to e if and only if *either*

(A) (1) c and e are causally connected, and
(2) there is some CCC of c and e which is not causally connected to c and is causally connected to e, and there is no CCC connected to c but not to e, or there is some quasi-CCC of c and e which is not quasi-causally connected to c and is quasi-causally connected to e, and there is no quasi-CCC quasi-causally connected to c but not to e, *or*

(B) c is causally connected to some event f and c is a direct condition of a causal connection between f and e, and f is causally prior to e.

This approach will allow us to handle cases involving single causes, if such cases are logically possible, without attributing causal efficacy or existence to non-occurrences. In addition, this change does not depart strongly from the original intuition underlying the analysis.

Finally, Bassham presents the following possibility: every event in the universe is causally connected to and counterfactually dependent upon a first cause c. It would seem that if c is causally prior to e, some event caused by c, then my analysis is unsatisfactory since there will be no CCC of c and e that is not causally connected to both c and e. In order to evaluate this case, two versions of it must be considered. Either this first cause is a composite event composed of sub-events or it is not a composite event. If the former, then, as I pointed out in my earlier treatment of causal asymmetry, we can treat that event c as causally prior to e in that the component events of the composite event will separately satisfy the definition of causal priority (ib. p. 773). If c is not a composite event, then this case can be treated as analogous to Bassham's 'single cause case'. The efficacy of c will depend in part upon certain negative events which are quasi-causally connected to e but not to c. Given such negative events, c will satisfy the conditions specified for causal priority to e.

Southern Methodist University,
Dallas, Texas 75275, U.S.A.

© DOUGLAS EHRING 1986

METAPHYSICAL REALISM AND HISTORY

By NORMAN MELCHERT

IN 'Metaphysical Realism' (ANALYSIS 42.1, January 1982, pp. 1-3) J. J. C. Smart attacked Putnam's restriction of belief to 'empirical realism' by providing an example designed to show that 'an ideal theory T1 could consistently be supposed to be false' (p. 1). The example invites us to imagine another theory T1*, like T1 in all empirically ascertainable ways, but positing that our four-dimensional space-time manifold is a cross section of a five-dimensional manifold containing another cross section having galaxies, planets, etc. T1 would be ideal for us in the sense that it would satisfy all possible observable and theoretical constraints. But it is quite conceivable that T1* should be the true theory. He concludes that metaphysical realism is true.

Peter Smith has challenged the argument involved (ANALYSIS 43.2, March 1983, pp. 74-8). He claims that Smart faces a dilemma. Either T1 is indeed empirically ideal, and then what does the 'embedding' of our universe in a five-dimensional manifold come to? Or T1* really does describe a causally unified whole, in which case T1 is not ideal. Even if, he goes on, the relations between the manifolds are merely geometrical and not causal, the most that can be said about T1 is that it is incomplete, not that it is false. But the conclusion needed to support metaphysical realism is that an operationally and theoretically ideal theory might be *false*.

These objections are powerful. Yet I think Smart's conclusion is correct. Rather than discuss Smith's objections in detail, I want to direct attention elsewhere. Arguments such as Smart's are usually conducted with the paradigm of physical theory in mind. What is envisaged is a theory made up of general propositions or laws which describe what always happens (under given circumstances). From such a theory predictions and retrodictions can be made which tie down ultimately to observation. The very notion of an approach to an ideal theory of this sort presupposes that what the theory is describing and explaining remains in some sense available. The evidence to which the theory is answerable does not, somehow, get lost or distorted irretrievably.

Such a physical theory, it is safe to say, will not tell us everything we may wish to know. In particular, for several well known reasons, it will not tell us about the actions of human beings in history. Knowledge of these is just as genuine as knowledge of the behaviour of the physical world. Yet in this sphere the evidence to which our 'stories' are responsible (stories here correspond to theories) is notoriously subject to loss and distortion.

Consider an experience that most of us have surely had. You are engaged in conversation and make a remark. It is taken ill, perhaps as an insult or slight. (Perhaps the remark, so interpreted, leads to a quarrel, to a war.) You did not intend it to be an insult, but you can see (after the fact), in the circumstances as they were (which you now become fully aware of for the first time), that it must have seemed to be such. In fact, did you not know better yourself, you would certainly draw the conclusion that it was an insult.

It may be possible to protest your innocence successfully. But there are circumstances in which protesting your innocence of any intention to insult will 'only make matters worse'. It will appear you are adding the crime of insincerity to the crime of insult. Nothing you can say or do will help. The best available evidence, and you can see this yourself, convicts you. And yet you are innocent. You might be able to convince others of that innocence if you could convince them that you were truly ignorant of the circumstances that seem to require the 'guilty' interpretation. Yet it may not be plausible that you *could* have been ignorant of them. (For all that, of course, you might have been.)

We must be careful here. It may well be that you are deceiving yourself, that the insult was 'Freudian', in the sense that it was not consciously willed but intended none the less. Sometimes we are justified in refusing to believe protestations of innocence. And yet who will say that it *must* be so? It seems perfectly intelligible that an innocent remark (which like all remarks must be interpreted in terms of the context of its utterance) may be mistakenly taken for an insult. And this on the very best evidence available.

There is no reason to believe that this kind of error can always be corrected subsequently either. Roughly, there might be two kinds of grounds on which correction could be made, first person grounds and third person grounds. Suppose that you, as honestly as you can, write your autobiography. In it you state what your intentions were and explain what caused them to be taken wrongly. You tell what you take to be the correct story about the incident. The response may be, 'Well, that's just what we would expect the guilty fellow to say, isn't it?' And this response, given all the relevant available evidence, may be the rational and justifiable one.

Suppose that some third person writes your biography; if the remark and its interpretation had serious consequences, note will have to be taken of it. And it may be that the best evidence available leads to the conclusion that you did intend to be insulting. Perhaps there is some possibility that your protestations are correct and that it was all a horrible mistake; but the overwhelming likelihood is that you are guilty. (If any philosophers are tempted to bite the bullet here and say that if the ideal story (the one which makes the best total sense of all the available evidence) describes the incident as an insult, then it *is* an insult, I simply ask them to put themselves into the place of a victim of such interpretation. Are they really willing to give up the authority for knowing their own intentions *completely*?)

Here we have a case where the 'ideal story' from the point of view of a community of historical enquirers may still be false. It differs from Smart's example in dealing with the interpretation of a particular event, a human action, and what to make of it. It points us to history rather than to physics. It is, moreover, grounded in an experience many of us have already had and does not depend upon positing an ideal physical theory which we may never get. Yet the conclusion is the same. The best evidence available, or *ever to be available*, may confirm a falsehood. And metaphysical realism is correct.

How does this argument compare with the one Smart offered? It does not have the appearance of completeness that is envisaged by Smart's T1. In a certain sense, T1 will 'talk about' everything that humans with their epistemological abilities can come to know. Historical stories will certainly not do that. But even now we know that T1 is complete only at a certain level or from a certain perspective. No one expects physics even to mention, let alone explain, the exploits of Napoleon or the poetry of Goethe – as such. I take it, however, that we humans have an ineliminable interest in description from what Dennett calls the intentional and personal stances. We cannot be content with physics alone. So my argument *supplements* Smart's argument in this respect.

Moreover, the conclusion of Smart's argument is that, though T1 is rational to believe, it 'might nevertheless be the case that T1* is true' (p. 3). This *possibility* of a disparity between the rationally ideal theory and the truth is, of course, all the metaphysical realist needs. Yet for historical facts, I suggest, we have more than just a possibility. In the case of circumstantial guilt, as well as in certain cases of successful lies and fortunate deceptions, we have good reason to believe that the ideal story not only *may* diverge from the truth but that it in fact *will* do so. For in these cases the relevant evidence is not ever-present for future historical enquirers to return to. In these cases the evidence is itself historical, composed of particular facts and circumstances, once occurring and then never again. And we have good reason to believe that even ideal historians of Peirce's future society of ideal enquirers will not be able to recover or reconstruct certain facts relevant to the interpretation of these historical events. In this respect, my argument is *stronger* than Smart's.

Finally, the historical argument is not subject to the apparently strong criticisms proffered by Peter Smith.

Lehigh University, Bethlehem,
Pennsylvania, U.S.A.
The University of Kent, Canterbury

© NORMAN MELCHERT 1986

A PRIORI PRINCIPLES AND ACTION-EXPLANATION

By PHILIP PETTIT

I NOMOTHETIC AND NON-NOMOTHETIC EXPLANATION

J. E. Tiles has recently tried to undermine a contrast which I wished to set up between action-explanation and the explanation of natural events [3]. I believe that my contrast can withstand his assault, though this is only clear when it is provided with deeper foundations than I supplied before. Here I try to indicate the sort of foundations available (see too [2]).

My earlier argument was that while the explanation of natural events is usually nomothetic, the explanation of action is never of this kind. I assumed that every event-explanation lends itself to generalization, gesturing at some principle or law; I mention the grounds for this assumption in the next section. A nomothetic explanation I defined as one in which the principle adumbrated 'is not itself so deeply espoused as to be held back from empirical questioning'. The principle must not be one of those 'truths which we would not see reason to give up, granted our other beliefs, under any imaginable circumstances'; it must not be 'contextually *a priori*' ([1], p. 95).

The reason I characterized action-explanation as non-nomothetic is that I took the principles which it presupposes all to be instances of a schema like the following: if someone desires something, and believes that by A-ing he can best get it, and other things are equal, then he A's. Such principles are bound to be contextually *a priori*. In any instance where one of them seemed to be falsified, the consequent failing to follow on the antecedent, we would naturally conclude that the antecedent could not have been realized after all.

Because action-explanation tries to bring behaviour under such copperfastened principles, I argued that it has a distinctive character. It consists in varying the antecedent attitudes ascribed to a person – subject to independent constraints – so as to represent him as an instantiator of the principles: that is, as a behaviourally rational agent. It takes its principles as unquestioned standards of rationality and tries to work out how agents must be attitudinally featured, given that they live up to such minimal norms.

The contrast which I alleged between the explanation of natural events and action-explanation is not that all natural explanations are nomothetic; they are not, since they may involve deeply embedded principles which enjoy contextually *a priori* status. It is only that some, perhaps most, are of a nomothetic character. Tiles acknowledges the qualified nature of this claim but finds other grounds for criticizing it. He argues that the distinction I invoke between nomothetic and non-nomothetic explanations is itself spurious.

In making his charge, he forecloses an important question of interpretation. The question is whether my definition of nomothetic explanation stipulates that the principles adumbrated there are always put forward as provisional and revisable; or whether it requires merely that they are always capable of being submitted to empirical questioning. Even from the summary of my claims given here, it ought to be clear that I need only defend the weaker thesis in order to sustain the contrast I want. But in the earlier statement of those claims I did not make this quite obvious; in some passages indeed I wrote as if I was committed to the stronger version of the contrast.

Tiles's charge is directed against this overstatement. He emphasizes, quite rightly, that at least 'in the everyday explanation of natural events the established laws of science are not called into question' ([3], p. 192). He concludes then that I did not after all make out the contrast which I was after.

At a first level it is clear what my response must be. This conclusion is not warranted because the criticism applies, not to any essential claim, only to an overstated version. The charge is off target.

But the response will hardly close the debate. Tiles may concede that there is the slighter nomothetic/non-nomothetic contrast between the two forms of explanation but deny that it is of any significance. If ordinary event-explanation does not question its principles any more than action-explanation, it is not obvious why the fact that those principles remain generally questionable should make such a difference.

II REGULARIZING AND NORMALIZING EXPLANATION

My aim is to show that there is independent ground for ascribing significance to the contrast: I will try to do this by presenting the distinction in the light of a deeper divide. This divide separates two ways in which an explanation may have force: by regularizing the explanandum or by normalizing it.

Every explanation of an event mentions a cause and mentions it in such a way as to gesture at an independently plausible principle: a principle linking the occurrence of the cause with the sort of event under explanation. That the explanans was satisfied does not tell us why the explanandum was realized unless there is a causal connection between the two. But the ascription of a causal connection is not sufficient on its own to explain. We ascribe such a connection but do not provide an explanation when we say that the event which the newspaper is going to report caused the tower to topple. What is also necessary for explanation is that we describe the events in such a way that we point to an independently plausible generalization about the co-occurrence of cause and effect. In the example on hand we would have needed to describe the cause, say, as a bolt of lightning.

Why is it that mention of a cause in a manner which gestures at a plausible principle has explanatory force? I believe that either of at least two different answers may be appropriate, depending on the case in question.

The standard account of where that force comes from is the regularizing one. This holds that to make an event intelligible is to succeed in presenting it as a particular instance of how the world works generally. It is to show that the event is unexceptional, being the routine operation of a recurrent process.

Fully articulated, the predicate 'explains' has more than the two places required for the explanans p and the explanandum q. In order to specify an appropriate generalization, it must leave room for mention of the possibilities by contrast with which p is explanatory and q is explained: say, o and r, respectively. And, more important for our present purposes, it also has to allow space for the assumptions against the background of which intelligibility is revealed: say, s. It is shorthand to say simply that p explains q. Properly speaking, we should say that p (rather than o) explains q (rather than r), given s (see [4], chapter 5).

The standard account of event explanation tells us something about the assumptions, s. On this account, what is assumed by way of background to explanation is that the world works to a causal order, embodying mechanical, if sometimes only probabilistic, regularities. The idea is that, given such a background, an event is rendered intelligible when we are shown that it is part of the dispensation involved.

The account is not implausible, and it certainly holds of some event explanations. It makes sense of the fact that the explanation of an event mentions a cause and gestures at an independently plausible causal principle. If it is assumed, as part of s, that the prevailing regularities include the sort of regularity to which the principle answers, then to show that the event is covered by the principle will be to regularize it.

On the regularizing account, explanation is marked by the role in which it casts the principle that it embodies. That principle has the status of a given, enunciating, or at least providing evidence for, a governing regularity. This observation is the cue for envisaging an alternative, non-standard account of event-explanation.

It might be that the principle represented a norm at which the world or some part of the world aimed, rather than just a datum about how it worked. In that case the explanandum would be made intelligible, not just by being shown to exemplify the world's regular mode of operation, but by being depicted as something that had to happen if the world was to continue to satisfy the principle that represents its norm. We can speak here of an explanation that normalizes the explanandum. The principle is a norm in the light of which the explanandum is required, rather than just a datum in the light of which it is regular.

The difference between regularizing and normalizing explanation comes in the assumptions against the background of which the account is offered. Suppose that we mention C in explanation of E, gesturing at a principle that links C-type and E-type events. The assumption under which E is regularized is that the C-E regularity is one aspect of the world's routine operation overall. The assumption under which it is normalized is that the regularity has a special sort of status.

The idea is that we know that the regularity is a goal for the system in question. More precisely, we know that the system is such that if E is necessary for the satisfaction of the C-E principle, then E occurs. Of course one ground for believing this would be offered by the knowledge that the C-E regularity actually obtains. But the idea is that we have independent reason for accepting it.

Suppose that conditions like the following obtain. First, there is a more or less well demarcated domain of explanatory concern, involving systems of a certain kind. Secondly, the systems in that domain are designed or selected or whatever so as more or less perfectly to fit a certain conception. And thirdly, we understand the conception sufficiently to be able to spell out norms at least some of which must apply to beings that fit the conception. The systems will satisfy the antecedents of some norms and in those cases they will generally satisfy the consequents too: generally rather than universally, because we can allow for blocks and imperfections, even ones we don't fully understand.

Suppose now that the C-E principle is one of the norms in which we articulate our understanding of such a system. In that case we will have independent reason for thinking that if E is necessary for the satisfaction of the principle – this, because C has occurred – then other things being equal, E occurs. The point will be clarified by illustration.

Imagine that we are presented with a robot which, allegedly, is designed to do housework. Suppose that we are convinced indeed that it does fit the conception adequately, being the product of reliable engineers. In such a case we will face an interesting explanatory task. As we watch the metallic creature pottering about the house, we will want to see how the various things it does – apart from those we come to regard as mere noise – are instances of housework. We will want to see how our conception of a houseworker is substantiated by the robot.

We can easily envisage examples of such explanation. We ponder about the robot's probing beneath the freezer until we realize that it is trying to extract some of the baby's toys. We puzzle over its puffing about the sofa until we see that it is doing its best to dust it. And so on.

The exercise is an example of normalizing explanation. Our conception of the houseworker gives us a vague cluster of norms such that we expect any domestic worker to satisfy the antecedents of some and, when it does, generally to satisfy the consequents too. Given that it registers that the sofa is dusty, we expect it to put the problem right. We substantiate our conception when we see various pieces of robotic behaviour as occurring in order that such norms should be fulfilled.

Less exotic illustrations of normalizing explanation are also available. Given the chess-playing computer programme, the pocket calculator or even a central heating system, it is clear that there is room and need for normalization. The exercise will substantiate our preconception of the system, relating its actions to antecedent conditions in a manner which shows those actions to be the sort of thing required of a system of that kind.

The idea of normalizing explanation will not be unfamiliar. The explanation envisaged corresponds to a widely endorsed conception of teleological explanation. According to that conception, an event E is teleologically explained as the means to a goal G, if it is caused to occur by the circumstance of an E-type event's being necessary for the goal. The conception fits our case, where the circumstance of an E-type event's being necessary for the obtaining of L – a circumstance independently describable just as the occurrence of a C-type event – causes E to occur also.

So much then for the two ways in which an event-explanation can have explanatory force. Although every such explanation mentions a cause and gestures at a principle, it may do so to either of at least two effects. Depending on the character of the relevant assumptions, the explanation may serve to normalize or merely regularize the event in question.

III ACTION-EXPLANATION IS NON-NOMOTHETIC, BECAUSE NORMALIZING

My claim at the end of the first section, maintained in face of Tiles's criticism, was that action-explanation was distinctively non-nomothetic. The question remaining however was whether this made for a significant thesis. It was not obvious, partly as a result of that criticism, that I had identified an important distinguishing feature. In this final section I hope to remedy that impression or lack of impression. I mean to show that its being non-nomothetic is explained by a distinct and certainly significant feature of action-explanation: specifically, that it is a consequence of the normalizing character of such explanation.

Consistently with being non-nomothetic, action-explanation just might be regularizing. It might be that the principles deployed were contextually *a priori*, although the purpose of the explanation was merely to represent action as part of the regular world. There may be some difficulties in substantiating the possibility but I know of none that is decisive.

Consistently with being non-nomothetic, action-explanation might equally be of a normalizing character. People are socially trained, no doubt on the basis of a suitable evolution and biology, to conform to a certain conception of the normal, rational agent. We understand this conception sufficiently well to be able to identify norms that we might expect then to apply. There is nothing incoherent in the idea that belief-desire principles are norms of this kind and that action-explanation is an enterprise of seeing how individual behaviour bears them out.

The two cases however are not on a par. If action-explanation is regularizing, it must be something of a surprise that its principles are *a priori*. Certainly there is no obvious reason why they should be of this standing. If action-explanation is normalizing on the other hand, it is only to be expected that the principles should have *a priori* status.

On the normalizing picture, the principles of action-explanation, relating beliefs and desires to behaviour, are norms which express our conception of the normal human agent. They are like the principles in which we might formulate our conception of the chess-playing programme, or pocket calculator, or central heating system. Understand what it is to be a normal agent and it will be undeniable, for each of these propositions, that anyone who satisfies the antecedent will, other things being equal, satisfy the consequent too. That is to say, the principles will be *a priori* knowable.

Given that we can account for its non-nomothetic character only on the normalizing picture of action-explanation, the argument to the best explanation suggests that we should indeed see it as normalizing. The upshot is that action-explanation is normalizing in the first place and, as a result of having this character, it is non-nomothetic too.

In the first section I offered an appropriate but apparently fragile response to Tiles's criticism. The position which we have now reached enables me to buttress that reply. It is an important feature of action-explanation that it is non-nomothetic. And this, even in the weaker sense which Tiles forces me to give to that characterization. It is an important feature because it is a significant one. Specifically, it signifies that action-explanation is an attempt to normalize behaviour, not just to regularize it.

Research School of Social Sciences,
Australian National University,
Canberra, ACT 2601, Australia

© PHILIP PETTIT 1986

REFERENCES

[1] Graham Macdonald and Philip Pettit, *Semantics and Social Science*, Routledge and Kegan Paul, London 1981.

[2] Philip Pettit, 'Broad-Minded Explanation and Psychology', in John McDowell and Philip Pettit, eds, *Subject, Thought and Context*, Oxford University Press, forthcoming.
[3] J. E. Tiles, 'Pettit on Revising Our Understanding of Individuals', *Analysis* 43.4, October 1983.
[4] Bas C. van Fraassen, *The Scientific Image*, Oxford University Press, 1980.

INTENTION AND PRACTICAL REASONING: A REPLY TO DONALD DAVIDSON

By CHRISTOPHER PEACOCKE

IN his Postscript to *Essays on Davidson: Actions and Events*, Donald Davidson reports that he has changed his attitude to writing public answers to criticism: he will no longer avoid it. Trying to achieve the beneficial effects noted by Davidson of writing a public reply, I here reply to Davidson's own criticisms of my views in that same volume.[1]

In 'Intention and Akrasia', I questioned Davidson's claim that the agent who does one thing rather than another and does so intentionally makes an outright judgement that the first, which he does, is better than the second, which he does not (EAE, pp. 23, 39, and 83–102). I also advocated a characterization of akratic action different from Davidson's.

Suppose an agent knows she could make a phone call which would be financially advantageous but personally unpleasant; akratically, she does not make the call. In the notation I offered, this case would be described thus. The agent judges that the fact that phoning in two minutes would be financially advantageous favours her so phoning:

> fav(I phone in 2 minutes, my phoning in 2 minutes is financially advantageous).

She also judges that

> fav(~(I phone in 2 minutes), my phoning then is unpleasant)

---

[1] See *Essays on Davidson: Actions and Events* (Oxford: O.U.P., 1985), edited by Bruce Vermazen and Merrill B. Hintikka. Henceforth I refer to this volume as EDAE. My 'Intention and Akrasia' is on pp. 51–73 of EDAE; Davidson's 'Reply to Christopher Peacocke' is on pp. 207–11. 'Intention and Akrasia' was discussing primarily Davidson's well-known paper 'How is Weakness of the Will Possible?', reprinted in his *Essays on Actions and Events* (Oxford: O.U.P., 1980), henceforth EAE. I have aimed to make the present reply relatively selfcontained: but one who has read all three papers will have a much better understanding of why I have concentrated on some points rather than others.

and that taking both reasons into account favours phoning:

> fav(I phone in 2 minutes, my phoning then is unpleasant but financially advantageous).

In this framework, her formation of the intention not to phone is akratic because she judges that both reasons together favour phoning; and neither her intention not to phone nor her acting upon it involve any outright judgement that not phoning is better than phoning.

In Davidson's framework, the case is described thus. For brevity, let us use some labels. We will be concerned with the content that prima facie *x* is better than *y* given that *x* is financially advantageous and *y* is not:

> R1 pf(*x* is better than *y*, *x* is financially advantageous and *y* is not).

It will also help to label

R2 pf(*x* is better than *y*, *x* is not unpleasant and *y* is)
r1 *a* is financially advantageous and *b* is not
r2 *a* is unpleasant and *b* is not.

Here *a* is a phoning in two minutes, *b* is an action which involves not phoning in two minutes. Davidson then characterizes our akratic agent as making four judgements. The first three judgements are of relativized prima facie contents, while the fourth is of an outright, unrelativized content:

pf(*a* is better than *b*, (R1 & r1))
pf(*b* is better than *a*, (R2 & r2))
pf(*a* is better than *b*, [(R1 & r1) & (R2 & r2)])
b is better than *a*.

That last content is judged in intentionally doing *b* rather than *a* on the Davidsonian account. The agent is akratic because that content is obviously contrary to the content which she judges prima facie to be the case relative to the conjunction of her reasons [(R1 & r1) & (R2 & r2)]. The principle of continence enjoins that if *a* is better than *b* relative to all the relevant considerations, then one should infer that *a* is better than *b* outright.

Davidson brings the following objections against my position. First, he says that my notation 'fav' mingles propositional attitudes with notions of inductive support, and 'the result seems to be a bad pun' (EDAE, p. 209). Second, while he would not mind giving up the word 'better', he holds that we need a general-purpose evaluative word if we are to represent the structure of practical reasoning (EDAE, p. 208); he also adds that for the same reason I will need '*some* term to help express the particular attitude which constitutes intention' (EDAE, p. 209). Third, he says that my claim that there is a failure of analogy between practical and theoretical reasoning in his treatment of the principle of continence stems from my following a different model from that which he advocates: I have not shown there is anything wrong with his way of doing it (EDAE, p. 208). In addressing these criticisms, I will argue that the issues are much less a matter of notation and more a matter of substance than Davidson implies; I will also argue that the issues are significantly connected.

In my notation, 'fav(I phone, *R*)' expresses the *content* of our agent's judgement when she judges that the reasons *R* favour her phoning. 'fav(I phone, *R*)' does not make a statement *about* her propositional attitudes. In particular then, it is not the right sort of sentence to feature in a theoretical inference, either on her part or on our part, to a conclusion about what she will do. Davidson represents me as wanting 'to infer from "fav(I ϕ, *E*) and *E* is the total available evidence" to "I will ϕ"' (EDAE, p. 209). But we ought not to describe what fills the second argument place of 'fav' as giving evidence: it rather gives reasons, reasons in favour of acting one way rather than another. I agree with Davidson that it would be an incoherent mingling both to maintain all I have done about 'fav' and to insist on the inference he represents me as wanting. I did not want that inference, but rather to maintain the following. The continent agent is the one who, when he judges a content 'fav(I ϕ, *R*) and *R* is the totality of relevant reasons', forms an intention with this content: that he ϕ's.

Am I denying the need for a general-purpose evaluative term in representing practical reasoning? In one sense I am on the contrary committed to the legitimacy of introducing such a term.[2] For I say that an agent may be judging something of the form 'fav(I ϕ, *R*)', whether the reasons *R* be moral, aesthetic, prudential or anything else. As far as I can see, I use this form in just the cases in which someone following a more Davidsonian construction would be willing to say that the agent judges that his ϕ-ing is, prima facie, good relative to the reasons *R*. So far, the issue does seem to be notational. Davidson suggests the highly appropriate term 'intention-worthy' as a general-purpose evaluative term (EDAE, p. 211). But can one judge that something is intention-worthy without yet forming an intention to do it? It is at this point that the issues become substantive and interconnected.

When I denied that in intentionally doing *a* rather than *b*, an agent is judging outright that *a* is better than *b*, I was not denying that there is some evaluative concept or other which the agent judges to apply to *a*. Indeed, to do so would arguably have been inconsistent with my insistence (in common with Davidson) that, even in akratic action, the agent has some reasons in favour of the course he chooses. What I was denying was rather this: that there is any evaluative concept, specific or unspecific, such that intending to ϕ can be identified with or at least involves the agent's judgement outright that his ϕ-ing falls under that evaluative concept. In the nature of identity, it has to be the same evaluative concept from occasion to occasion for this claim to be true. The claim that there is any such concept is not established just by arguments which demonstrate the need for some general-purpose evaluative concept in representing practical reasoning. If there is such a concept, then, for example, it is one about which the agent has to make different and incompatible judgements in two otherwise similar situations in one of which he acts akratically and in the other of which he does not. Let us call the claim that there is an evaluative concept such that intending to ϕ necessarily involves the outright judgement on the agent's part that his ϕ-ing falls under that concept 'the evaluative-judgement theory'. It is a theory about the nature of intention.

---

[2] This sentence corrects the second paragraph of p. 58 of 'Intention and Akrasia', which makes the mistake of supposing that if a theory does not explicitly use a general-purpose evaluative concept, it is not committed to its legitimacy. This is an error, since a theory may define the concept implicitly – albeit perhaps in a way of no great philosophical interest.

When we look at the full range of intentional action, akratic and continent, it seems to me hard to defend the evaluative-judgement theory. But instead of elaborating the arguments I gave against it (EDAE, pp. 55–6), I want here to consider the consequences of rejecting it. On Davidson's theory, the *same* evaluative notion features both in the content specification which fills the first place in the 'pf' connective and in the outright content judgement of which is involved in forming an intention. This remains true if some evaluative notion other than 'better' is used. But if we reject the evaluative-judgement theory of intention, there will not be anything which can both be an evaluative content featuring in the first argument-place of 'pf' and be such that outright judgement of it is distinctive of intention. It becomes plausible that the only content an attitude to which is distinctive of intention is the content of the intention itself. If that is so, then when we are representing practical reasoning, it will be necessary to indicate for which contents judgement is in question, and for which intention is in question. That was what I did in my formulation of the principle of continence above.

A second consequence of rejecting the evaluative-judgement claim is for the parallelism with inductive inference and Carnap's requirement of total evidence. Davidson rightly says that there are many versions of a principle of acceptance which are concerned with total evidence (EDAE, p. 208). However the parallel Davidson proposed is incompatible with rejection of the evaluative-judgement theory. Davidson's principle of continence says

> From 'pf(*A* is better than *B*, *R*) and *R* is the total available evidence' infer '*A* is better than *B*'.

This principle should not be disputed taken neat. But it can only be taken as the principle violated by the akrates if we accept in addition that forming an intention to do *A* rather than *B* must involve judging that *A* is better than *B* outright. We cannot accept that claim about forming an intention if we are rejecting the evaluative-judgement theory. In 'Intention and Akrasia' I set up a different parallel for two reasons: by that stage in the paper I had already argued against the evaluative-judgement theory, but I wanted to preserve Davidson's important insight that some apparatus of relativization is required if we are to characterize akrasia properly.

I have been arguing that there is a substantive, unconfused alternative to Davidson's account which still preserves many of the most attractive features of his account, including some parallels between inductive and practical reasoning. This alternative is one on which the irrationality of the akrates is confined solely to irrationality in action and the formation of intentions: he need not be making irrational judgements at any level. In choosing between Davidson's theory and this alternative, we have to concentrate not on their common parts but on their differences. These differences flow ultimately from different stances on the evaluative-judgement theory of intention. But it would be a pity if a critic, finding the evaluative-judgement theory implausible and failing to realize the possibility of an alternative distinct from Davidson's, thought himself forced thereby to reject all the Davidsonian features still present in the alternative.

King's College (KQC),
London WC2R 2LS

© CHRISTOPHER PEACOCKE 1986

THE VOLUNTARINESS OF BELIEF

By JAMES MONTMARQUET

BELIEF, unlike action, is not – not directly – under our voluntary control. Call this the Asymmetry Thesis: that belief and action are 'asymmetrical' with respect to direct voluntary control. There is a strong temptation to think that this, or some slightly qualified version of it, is true. The argument of this note, however, is to the opposite effect. Briefly, it is that reasons for action play a role in the determination of action which is analogous to the role played by reasons for thinking-true in the determination of beliefs; and that, consequently, if the controlling influence of reasons on actions is compatible (as I assume it is) with the voluntariness of action, the same is true with respect to the influence of reasons for thinking-true on beliefs.

In his essay, 'Deciding to Believe', Bernard Williams puts the case very strongly for the absence of (direct) voluntary control of belief:

> If I could acquire a belief at will, I could acquire it whether it was true or not; moverover I would know that I could acquire it whether it was true or not. If in full consciousness I could will to acquire a 'belief' irrespective of its truth, it is unclear that before the event I could seriously think of it as a belief, i.e. as something purporting to represent reality (*Problems of the Self*, Cambridge, 1973, p. 148).

Now some points of detail in Williams's argument may be challenged;[1] but I shall accept it as it stands – and raise this question. Could we construct an analogous argument purporting to show, by way of a *reductio*, that action, too, must be regarded as independent of the will?

Williams's view is that one cannot simply bring about a belief in onself, if this means independently of what one has reason for thinking-true. But, analogously, can one simply initiate an action, say, a hand movement, independently of what one takes oneself to have reason *to do*? Let me offer six comments in support of the view that action and belief are alike in this regard.

(1) It may be urged that action is far less severely constrained by reasons: that whereas I can do whatever I regard myself as having some reason to do (as long as this is within my current physical power), I can only believe what I take myself to have *adequate* (or sufficient) reason for thinking-true. Reply: even if we assume that the latter is really true about belief, and I shall question this in a moment, it is still unclear that there is a genuine source of asymmetry here. Thus suppose that John believes that he has *some* reason to do X, but much better reason not to do it; and that John takes himself to have some reason to believe that p, but much better reason not to believe it. *Given* all this, is John any less constrained in what he will (or can) do than what he will or can believe? Intuitively, it is far from clear that he is.

(2) It may be urged that, under the circumstances just specified, it remains true that John can do the act he regards less well because there is always the possibility that he will come to exhibit *akrasia*, weakness of the will. Reply: this suggestion is mistaken for two important reasons. First, it seems quite plausible to suppose that *akrasia* besets theoretical as well as practical reason. Isn't it quite easy to imagine cases of persons who recognize that, given the evidence, they should believe that p (that they have sufficient reason to believe that p), but who still cling for epistemically disrespectable motives to the opposite belief?[2] If this were not so, it would be hard to understand all of the recent attention paid by

---

[1] See (e.g.) Barbara Winters' 'Believing at Will', *The Journal of Philosophy*, 76.5, pp. 243–55.

[2] For an extensive discussion and defence of the existence of such incontinence, see Alfred Mele, 'Incontinent Believing', *Philosophical Quarterly* (forthcoming).

epistemologists to the notion of 'epistemic responsibility'. Second, and more fundamentally, it is a deep mistake to think that the phenomenon of *akrasia* – practical or theoretical – in any way counts toward voluntariness. As the very name, 'weakness of the will', suggests, this phenomenon marks a certain kind of *limitation* on one's active powers; its possibility is hardly a necessary condition of voluntary agency.

(3) It may be urged that the impossibility (if such it be) of doing what one does not think that one has adequate reason to do – that this is somehow a *logical* rather than a causal fact; that it simply follows from the fact that I intentionally raise my arm, that I believed that there was a sufficient reason to do this. The reply to this is obvious. It seems no less, and no more, certain in the case of belief that the analogous implication holds: if I believe that *p*, I believe that I have adequate reason to believe that *p*. If such an implication is compatible in the latter case with the causal determination of beliefs by reasons, it should be equally so in the former.

(4) In a subtler and more limited vein the following may be urged. Given my beliefs about what I have reason to do at any given time, there will be various actions (if we focus on the physical, various ways of moving my body) which equally well realize my ends. The doxastic sphere, however, presents no such indeterminacy. Given what I have reason for thinking-true, I am not similarly 'free' to choose among incompatible possibilities; rather, whenever I think that there are equally good reasons for believing *p* and believing not-*p*, I am constrained to believe *neither*, and to suspend judgment. Reply: of course this asserts only a very restricted and seemingly not very fundamental asymmetry between belief and action (since it limits itself to the case of choosing among options thought to be equally acceptable), but even so it oversimplifies matters. For there are two very different cases of epistemic indifference to be considered. The first is the above one in which there are equally good reasons for believing *p* and for believing not-*p*; the second is where there are equally good reasons for believing *p* and for not believing *p*. In the first of these, epistemic rationality bids one to suspend judgment; in the second, it yields no determinate answer. *Each* of these has possible analogues in the case of action. I may have good reason to do *A*, equally good reason to do some incompatible act *B* and better reason to refrain from doing either; likewise, I may have good reason to do *A* and equally good reason not to do *A*.

(5) It may be urged that there is this difference between actions and beliefs governing their relations to wanting:

> If I want to do *X*, it follows that I at least have some reason to do it; but if I want to believe *X*, no such reason to believe *X* ensues.

Now the effect of this, the argument would go, is that it is often the case, background considerations being what they are, that if I merely want to do *X*, I have not only some reason, I have a perfectly sufficient reason, to do *X*. In such cases, then, the argument concludes, one's actions are responsive to one's desires in a way that beliefs can never be. Reply: call such acts – where background circumstances and the type of act in question are such that wanting is a sufficient reason for doing – call them 'gratuitous acts'. Now the correct response to this is to point out that while there are indeed gratuitous acts but not gratuitous beliefs, since many (perhaps most) voluntary acts are *not* gratuitous, this by itself proves nothing with respect to the greater voluntariness of action. For the determinants of the voluntariness and the (possible) gratuitousness of an action are evidently not the same. Broadly speaking, we may say that the determinants of the former pertain to the sources of an agent's reasons *themselves* – especially to the absence of certain kinds of external intervention and internal compulsion; these factors have clear analogues in the case of belief: various forms of 'mind control', etc. By contrast, the determinants of gratuitousness have to do with the source of the *sufficiency* of one's reasons: that such sufficiency be grounded merely in the having of certain wants. The fact, though, that belief and action are assymetrical with respect to gratuitousness has no tendency to show that they are asymmetrical with respect to voluntariness.

(6) Finally, G. W. Barnes ('The Conclusion of Practical Reasoning', ANALYSIS 43.4, October 1983) has argued that the determination of beliefs and actions by reasons are dissimilar in this respect. A given belief may be unreasonable (i.e. unwarranted by the canons of deduction or induction) *simply* in the light of the agent's reasons for it and regardless of whether any other states of affairs obtain or fail to obtain; but, Barnes maintains, a reason for action can never be ('internally') unreasonable in this way: whether or not an act is reasonable in the light of a set of reasons will always depend on further circumstances (see esp. 195–6). The consideration underlying Barnes's contrast seems to be this: an agent's reason for doing *A* (in the sorts of cases, at any rate, in which he is interested) will always include some belief to the effect that by doing *A* the agent will promote some further end he has. And it is this belief (together with an appropriate desire for the end in question) which always confers a *certain* reasonableness on the agent's act – albeit one which may be overthrown by further considerations (in which case the act will be 'externally' but still not internally unreasonable).

However, even if Barnes is correct about such a dissimilarity – and the existence of known fallacies of deontic logic (reasoning) would surely indicate that his distinction marks off *instrumental* (means-end) practical reasoning from other sorts, *including* non-instrumental deontic arguments[3] – this issue of voluntariness remains substantially untouched. Consider the following analogy. A being capable of drawing logical inferences, given appropriate motivation and knowledge, will come to be constrained to some extent in his thinking by rules of logic. But is his thinking, all things considered, more constrained (therefore in some sense less voluntary) than a being who lacks this capacity? Well, even at its greatest restrictiveness, this constraint serves to govern a capacity which the other being quite *lacks*. Surely, then, the thinking of the being capable of logical inferences, while subject to an additional constraint, is not *less* voluntary than the other's. If anything – since it seems that voluntariness would be augmented by such a capacity – this being's thinking would be more voluntary, i.e. more subject to voluntary control. Likewise, it seems to me, insofar as theoretical reasoning involves the operation of a constrained capacity for logical inference not present in its practical counterpart, this would not make the formation of beliefs any less voluntary than the performance of actions – if anything, the very opposite.

A final point. We have been assuming thus far that the Asymmetry Thesis stands or falls according to certain analogies or disanalogies governing *reasons* – for belief and for action. But it may be questioned whether Williams's own argument, which does not explicitly mention reasons of either sort, turns on this consideration at all. His argument, it may be said, depends merely on the incompatibility of these two things: my thinking of *X* as a belief of mine; my arriving at *X* in 'full consciousness' that I am adopting it 'irrespective of its truth'. But consider what the latter involves. If I do arrive at a belief with such a consciousness, then surely I will *also* have adopted it for reasons other than ones I take to be reasons for thinking it true. Hence any difficulty we may have reconciling the former with the voluntariness of belief will also carry over to the latter. (Indeed, what distinction can there be between arriving at a belief because one thinks that something is true and arriving at a belief because one thinks that one has adequate reason to think that it is true?)

Washington State University,
Pullman, Washington 99164-5130, U.S.A.

© JAMES MONTMARQUET 1986

---

[3] The clearest case of what would appear to fit Barnes's description of an 'internally unreasonable' action would be one based on the following fallacious deontic argument: 'I ought to do *p* only if *q*; *q*; therefore, I ought to do *p*'. I am indebted to Harry Silverstein for discussion of this point.

REMORSE AND REGRET: A REPLY TO PHILLIPS AND PRICE

By L. N. Zoch

In their article 'Remorse Without Repudiation' (Analysis 28.1, October 1967) D. Z. Phillips and H. S. Price argue that an individual may experience remorse even when he believes his past action to have been morally justified, using the following example.

> A possessive mother opposes her son's marriage. The son is in no doubt that his mother's possessiveness is mistaken, but he does not want to hurt her. She has sacrificed and suffered much for his sake. He knows that if he marries she will go to pieces. He asks himself what he ought to do. Eventually, he decides to marry, but he feels remorse for what he has done to his mother.

The question is whether it can really be *remorse* the son feels for what he has done to his mother, given that he 'is in no doubt that his mother's possessiveness is mistaken', i.e. given his unwillingness to repudiate his decision to marry. I shall argue that it cannot be, at least I shall argue that Phillips and Price have not shown that we ought to think that it can be.

Phillips and Price claim that to deny that the son might feel remorse here is to commit oneself to a view of 'the relation between what a person decides to do in a situation, and good and evil' which in fact identifies 'good with what ought to be done and evil with what ought not to be done'. They claim that this is a silly view, and claim, furthermore, that in holding such a view one would be denying moral possibilities. Our authors are right on these last two points, but wrong on the first. As the above view does not allow for the possibility that in acting rightly one may yet do some evil, it certainly is a silly view. And as this view implies that our moral duty extends only to acting rightly in the first place and no further – that 'Once the decision has been made there is no need to worry about the other alternatives' – it certainly obscures moral possibilities. But plainly one does not have to hold such a view in order to deny that the son might feel remorse. It is not as if we can only understand the son to have harmed his mother and to be fully conscious of this fact if we picture him remorseful, as Price and Phillips would like us to do. Different pictures of the son are possible here, and, given our central assumption about the son, I would suggest that a different picture is required.

Phillips and Price offer us two conflicting (and extreme) pictures of the-son-after-his-decision-to-marry and ask us to choose between them. We do not have to choose between them. If we did, we should certainly have to concur with the choice of our authors, for consider the alternative we are given. This is a picture of the-son-after-his-decision-to-marry as feeling *regret.*

> The man who felt regret about harming his mother might speak like this: 'The last thing I wanted to do was to hurt the old lady, but it was unavoidable. She brought it on herself. These things happen, and we have to make the best of them.'

As Price and Phillips coolly note, this picture fails completely to reveal 'the *depth* of the problem' that might exist for a son who has harmed his mother in the way that the son of our example may be supposed to have harmed his mother. But consider a different picture – a picture of the son feeling not regret but *remorse.*

> The man who feels remorse cares deeply because he, a son, has harmed his mother. He is not blind to the extent to which his mother has brought the situation on herself, but, nevertheless, he feels shame and sorrow for the harm he has caused her and feels responsible for it in a way which the man who feels mere regret does not.

Here (it is suggested) is language better suited to the possibilities of our example. Equipped with this kind of language, the language of remorse, we could readily depict 'a son deeply concerned at having harmed his mother'. We could not do this equipped only with the language of regret.

This is nonsense. I repeat: we don't have to choose between these two pictures. And we *can* form an adequate picture of the-son-after-his-decision-to-marry using the notion of regret. It is a marvellous account of the concept of regret that is given by Phillips and Price. We may wonder, to begin with, whether a man who spoke as our man above speaks – 'She brought it on herself. These things happen . . .' – did even *regret* having harmed his mother. At all events, it is clear that a man who felt regret about having harmed his mother *might not* speak as our man above speaks and may indeed speak in quite different tones, and that the language of regret *might* therefore suffice to represent a son 'deeply' concerned about having harmed his mother. I submit that it would suffice.

Although I would insist that 'shame' be left out of the story (it being too close a cousin of 'remorse'), the 'remorse' postulated by Phillips and Price tells us nothing that 'regret' couldn't tell us, if its likely nature in such circumstances as we are considering was adequately conceived. *Of course*, if it is assumed that if a person feels regret then he or she feels the sort of regret our man above feels, it will not be possible to reveal the depth of the problem that may exist for that person. But there is absolutely no reason to make this assumption. The concept of regret, as Roger Trigg notes in his paper 'Moral Conflict' (*Mind* LXXX, 1971), does tend to carry 'a rather mild flavour about it', and 'The word does not do justice to the horror which we might feel at some of the things we have to do'. Yet, as Trigg also notes, despite this 'the basic notion of a reluctance to do something which we see as being bad in some respect still holds'; and no matter how great the agony, if there is *no self-condemnation*, it will be inappropriate to introduce the concept of remorse. True, ideas of sorrow, pity, compassion and the like go very naturally with the idea of remorse, but, if one stops to think about it, they go just as smoothly with the idea of regret (though not with the Phillips-Price brand of regret). For clearly regret as well as remorse *can* be (e.g.) deep or not so deep, huge or minor, lasting or passing, crippling or just 'a worry'; and to talk of 'mere' regret as Phillips and Price do almost looks like a trick. Again, it is true that regret is not a *moral* concept in the way that remorse is. But this does not mean that moral possibilities are precluded through our thinking of certain types of situations in terms of regret rather than remorse, and to think that it does mean this is simply a mistake. One might as well say that in denying that a piano composition is a 'sonata' for the reason that it lacks a particular sort of structure one is denying a musical possibility.

Unfortunately, Phillips and Price make this mistake. This leads them into utter confusion. They write:

> What is meant if someone says that a man cannot feel remorse in the situation we have described? Is a man simply misusing language if he says that he feels remorse about the harm he has done without repudiating the decision which led to it? But the question at issue is not a linguistic one, but a *moral* one. When someone says that a person cannot feel remorse in the way we have suggested, he is, perhaps without knowing it, speaking for himself. It is as if he were saying, 'I cannot see it in that way. I do not see how I or anyone else could feel remorse in that situation'. But it is a big step from a person's inability to see remorse here, to the general assertion, 'It *cannot* be seen in that way'.

There is no 'big step' here, nor is there an 'inability', only a philosophical answer to a philosophical question – and the accusation of moral blindness we can ignore. To use our authors' preferred style, it is decidedly *not* as if he (I) were saying 'I cannot see it in that way. I do not see how I or anyone else could feel remorse in that situation'. He is (I am) saying rather, 'I see what you mean, I see what you're getting at all right. I see the moral possibility you have in mind. Only you don't need to stretch the concept of remorse to cover that possibility; given a suitable moral charge 'regret' will do the job. Save 'remorse' for those cases where an agent sees himself as having acted improperly'. He is (I am) stating a straightforward philosophical thesis: that a certain (as it happens moral) concept, 'remorse', fails of application in a certain type of situation – that characterized by an agent's unwillingness to repudiate a past action even though it has given rise to evil.

D. Z. Phillips has had occasion to reaffirm the thesis he and Price recommended in this journal. On page 38 of his recent book *Through A Darkening Glass: Philosophy, Literature and Cultural Change* (Blackwell, 1982), in an essay entitled 'Some Limits to Moral Endeavour', Phillips writes, 'as I have argued elsewhere, even after a person has decided what he must do in [situations of moral dilemma] he may still feel remorse for having committed the evil which his decision inevitably involved'. Phillips and Price fear that if we deny this we must deny the reality of moral dilemmas. But the argument of this essay does not commit us to the view that 'The moral house can always be put in order' (as Phillips puts it on page 47 of his book). To reject a proffered description of a possibility is not to reject the possibility; and our present conclusion is quite consistent with the thought that 'very often there is no clear choice between good and evil'.

University of Melbourne,
Parkville, Victoria 3052,
Australia

© L. N. Zoch 1986

PIKE'S OCKHAMISM

By John Martin Fischer

I

NELSON Pike has presented an argument which purports to establish the incompatibility of God's omniscience and human freedom ([6], [7]). Elsewhere, I have defended this sort of argument against a compatibilist strategy which I dub 'Ockhamism' ([2]; see also [1]). Recently, Pike has 'switched sides' to argue that my defence of the incompatibilist argument overlooks the force of the Ockhamist criticism of it ([8]). I intend here to set out one version of the incompatibilist argument and explain Pike's recent answer to it (on behalf of the Ockhamist). In another paper, I have argued that the Ockhamism recently suggested by Pike is unacceptable insofar as it renders God's existence unduly dependent on *human action* ([3]). In this paper, I shall supplement my criticism of Pike's approach by arguing that Pike's Ockhamism entails that God's existence is dependent, in an unsuitable way, on *God's own actions.* I believe that the two strands of criticism of Pike's defence of Ockhamism provide an entirely convincing refutation of it.

II

Let me begin by setting out one version of the argument for incompatibilism. I shall assume here, following Pike, that 'God' is a 'title-term' or 'role-indicator' – a non-rigid designator whose descriptive content specifies a certain position or role. Thus, to say that God exists is to say that there exists someone who fills the role of God. In contrast, 'Yahweh' is taken to be a proper name. Further, I assume that among the essential properties of God are omniscience and eternality (here understood as 'sempiternality' or 'everlastingness'). One can distinguish between two conceptions of God's essential attributes. On one conception 'God is essentially omniscient' is interpreted as a *de dicto* attribution: necessarily, whoever fills the role of God is omniscient. On a contrasting conception the attribution of essential properties is *de re*: if Yahweh is God, omniscience is essential to the personal identity of Yahweh. The conception of God's essential attributes which I shall adopt here is the *first* one (for discussions of the second conception, see [8] and [3]).

Crucial to the incompatibilist argument is the 'fixed past constraint' on power attributions:

(FPC) It is never in any person's power at a time T so to act that the past relative to T would have been different from what it actually was.

I shall now present the incompatibilist argument (I follow my presentation of the argument in [2] and [3]). Suppose that Jones does X at T2 and that Yahweh is God. It follows from Yahweh's omniscience and eternality that He believed at T1 that Jones would do X at T2. Now if Jones is free at T2 to refrain from doing X, then:

(1) It is within Jones's power at T2 so to act that Yahweh would have been God at T1 and would have held a false belief at T1, or
(2) It is within Jones's power at T2 so to act that Yahweh would have been God but wouldn't have held the belief He held at T1, or
(3a) It is within Jones's power at T2 so to act that Yahweh would not have existed at T1, or
(3b) It is within Jones's power at T2 so to act that Yahweh (though existing at T1) would not have been God at T1.

(1) is ruled out by God's essential omniscience, and (2), (3a), and (3b) are ruled out by (FPC). If sound, the argument can be generalized to show that God's existence is incompatible with human freedom to do otherwise.

The Ockhamist wishes to resist the argument by pointing out that (FPC) applies only to a sub-class of facts. There does seem to be a distinction, even if we cannot formulate it precisely, between hard and soft facts about times. A hard fact about a time is a 'genuine', non-relational fact; a soft fact about a time is 'relational' – it is not only about the time, but also about another time or times. Both the incompatibilist and the Ockhamist agree that there is such a distinction. Further, both agree that (FPC) applies to all hard facts about past times; that is, both agree that one cannot act in such a way that a *hard* fact – a non-relational fact – about some past time wouldn't have been a fact. And whereas the hard facts about the past are, in this sense, fixed, both parties to the debate agree that some (though not all) soft facts about the past are not fixed (for examples of soft facts about the past which are, nevertheless, 'fixed', see [9], p. 165, and [4], pp. 432–3).

An Ockhamist, then, may wish to block the incompatibilist's argument by employing the hard/soft fact distinction to defend (2), (3a), or (3b). The Ockhamist is a compatibilist – he believes that it is possible that both 'God exists' and 'Humans are free' be true. In 'Freedom and Foreknowledge', I offered an argument that Yahweh's belief at T1 should be construed as a hard fact about T1, and thus that it is not open to the Ockhamist to defend (2) ([2], pp. 76–8). Pike does not dispute this point, and I shall take it as given, in this discussion. Also, on the *de dicto* interpretation of God's essential attributes, Pike provides no reason to doubt that the existence of Yahweh at T1 is a hard fact about T1, and so it doesn't seem to be open to the Ockhamist to defend (3a). But Pike correctly points out that I have offered no compelling reason to think that the fact that Yahweh fills the role of God at T1 is a hard fact about T1; thus, it might remain open to the Ockhamist to deny (3b). That is, Pike insists that it is open to the Ockhamist to say that a person can have it in his power so to act that the individual who actually is God *wouldn't* be (or have been) God.

III

I shall now develop an argument that if (FPC) does indeed rule out (2) – that is, *if* Yahweh's belief at T1 is a hard fact about T1 and (FPC) is true – then God's existence will be unacceptably dependent on *His own actions.* I shall argue that, relative to certain natural and plausible assumptions, the claim 'God exists', is *necessarily false*, insofar as it is admitted that God's prior belief is a fixed fact about the past. Thus, it will be seen that Pike's defence of Ockhamism is untenable.

I have already said that God is essentially eternal, i.e. that God is eternal in every possible world in which He exists. By this it is meant that, for every world w in which God exists and time T, 'God exists' is true at T in w. On the conception of 'God' adopted here, different individuals occupy the role of God in different logically possible worlds. Thus far I have also left it open that, although 'God exists' is true at each time in a given possible world w, there exist different persons who occupy the role of God at different times in w. But I shall adopt the assumption that, if a person fills the role of God in a possible world at time T, then that *same* individual occupies the role of God at all times in that world. So, whereas there are different individuals who are God in different possible worlds, there are *not* different individuals occupying the role of God in the same possible world. This reasonable assumption rules out 'God-switching' in a possible world.

Also, I shall assume that God has the essential property of omnipotence. Of course, it is notoriously difficult to give an acceptable characterization of omnipotence. But for the purpose of this discussion it will not be necessary to produce such an account. I shall simply assume that, if an individual is omnipotent, then he can perform an act such as causing the Statue of Liberty to fall. That is, I shall take it that, on any adequate definition of omnipotence, it will follow that an omnipotent agent can cause the Statue of Liberty to fall.

It follows from the assumption that God has the property of essential omnipotence that if an individual (say, Yahweh) is God in this world, then He is omnipotent in this world. Further, it might be thought (and Pike has argued that it is so) that there are *certain* acts which Yahweh can do but are such that, if He were to do them, Yahweh would not occupy the role of God ([5]). So, even if Yahweh actually occupies the role of God, He *can* sin. (For instance, He has the power to torture an innocent baby.) Of course, if Yahweh is actually God (and thus, on any reasonable conception of God, perfectly good), we know that he *won't* sin; but nevertheless, He has the power to do so, and if He were to exercise this power, then He *wouldn't be* (or ever have been) God.

So there are certain acts which are such that, if Yahweh were to exercise his omnipotence and perform them, He wouldn't occupy the role of God. But I claim that our conception of God requires that there be at least certain acts (other than the one God actually performs) which are such that, if Yahweh were to perform them, He would still be God. That is, my claim is that a reasonable conception of God requires not just that the individual who occupies the role of God be omnipotent; it also requires that the individual who is God can *exercise* His omnipotence (by doing something other than what He actually does) *and* still be God. Whereas if Yahweh is God, Yahweh's doing other than what He actually does shouldn't *entail* that Yahweh is still God, it should at least *permit* it – I think that we should reject a conception of God according to which God's doing other than He actually does entails 'role-abdication'. Let us then accept the assumption that God is not 'essentially role-abdicating':

(NRA) For any time T, there is at least *one* action other than what the person who is God actually does which is such that, if He were to do it at T, He would still be God at T.

I shall now present an argument which mimics the incompatibilist argument set forth above and which shows that Pike's defence of Ockhamism is untenable: if God's belief is a fixed fact, then 'God exists' is necessarily false, relative to the quite attractive assumptions about God developed above. Suppose that the Statue of Liberty does not fall at T2 and that God exists. It follows that Yahweh (being God) believed at T1 that it wouldn't fall at T2. If Yahweh has it in his power at T2 to cause the Statue of Liberty to fall at T2, then:

(1′) It is within Yahweh's power at T2 so to act that Yahweh would have been God at T1 and would have held a false belief at T1, or
(2′) It is within Yahweh's power at T2 so to act that He would have been God but wouldn't have held the belief He held at T1, or
(3a′) It is within Yahweh's power at T2 so to act that He would not have existed at T1, or
(3b′) It is within Yahweh's power at T2 so to act that Yahweh (though existing at T1) would not have been God at T1.

(1′) is ruled out by God's essential omniscence. And (2′) and (3a′) are ruled out by (FPC); at least, for the purposes of this discussion, we shall assume this to be so. (Pike grants that Yahweh's belief at T1 is a hard fact about T1. And insofar as (FPC) applies to divine persons as well as human persons – as is eminently plausible – (2′) must be ruled out. The only way to avoid this conclusion, given that Yahweh's belief is considered hard, is to say that Yahweh, being God, can affect *hard* facts about the past, in a way in which mere humans can't. But this position seems unacceptable. If a fact is genuinely a *hard* fact about the past, how can a temporal God now 'alter' it?) Thus, the only way in which we can allow Yahweh the freedom to cause the Statue of Liberty to fall at T2 would be to affirm (3b′). On this approach, if Yahweh were to cause the Statue of Liberty to fall at T2, then He would have had a false belief at T1 (the belief that it would not fall at T2) and thus wouldn't have been God at T1. (If omniscience were *essential* to the personal identity of Yahweh (contrary to the assumption adopted in this paper), then we obviously could not accept (3b′). I develop this sort of criticism, adopting the *de re* conception of God's essential attributes, in [3].) And in virtue of the assumption which rules out 'God-switching', if Yahweh were to cause the Statue of Liberty to fall at T2, then He would not be God at T2. But note that the argument is perfectly *general* – it applies no matter what action we pick. So it follows that, for *any* action X (other than the action which Yahweh actually performs), if He were to do X at T2, then He wouldn't be God at T2. But this violates (NRA), the assumption that the individual who is God is not essentially role-abdicating. If we accept (3b′) in the case of causing the Statue of Liberty to fall, we must accept it for *any* action; but to do so would violate (NRA). Thus, (3b′)

must be ruled out. And if so, it follows that Yahweh *can't* cause the Statue of Liberty to fall at T2, and thus is not omnipotent. But this clearly contradicts our assumption that God *is* omnipotent.

The argument presented above shows that, relative to the assumption that God's exercise of his freedom to do otherwise should not require role-abdication, it is necessarily false that any individual occupy the role of God, where the role includes essential eternality, omniscience, and omnipotence (interpreted as the properties are interpreted above) and where God's prior belief is taken to be fixed. The assumption (NRA) is surely plausible: whereas it is reasonable to assert that Yahweh wouldn't be God if He were to sin, it is not reasonable to assert that He wouldn't be God if He were to do *anything* other that what He actually does. While it might strain ordinary intuition to claim that it is *possible* that the individual who is God abdicate his role by doing other than what He actually does, it surely is unacceptable to claim that it is *necessary* that the individual who is God abdicate His role if he exercises His freedom to do otherwise. And if 'God exists' is necessarily false, then it follows trivially that God's existence is *not* compatible with human freedom, and Pike's Ockhamism is unacceptable.

IV

In this paper I have shown that, if one accepts (as Pike does) that God's prior belief is a fixed fact about the past, then the identity of God would depend inappropriately on God's *own* actions. This argument complements an argument which I present elsewhere that, on the approach to God's essential attributes adopted here (the *de dicto* approach), God's identity would depend unduly on human actions. Together, the two strands of argument lend considerable weight to the claim that, if God's prior belief is fixed, then Ockhamism is untenable, on the *de dicto* conception of God's essential attributes. In [3] I argue that, if God's prior belief is fixed, then Ockhamism is also untenable, on the *de re* conception of God's essential attributes. Thus, an Ockhamist will be forced to *deny* that God's prior belief is a *hard* fact about a past time; to do so, I believe that he must confront my argument that God's belief *is* a hard fact ([2]). If God's belief is a 'genuine', 'non-relational' fact, then Ockhamism must be rejected.[1]

Yale University,
P.O. Box 3650 Yale Station,
New Haven, Connecticut 06520, U.S.A.

© JOHN MARTIN FISCHER 1986

---

[1] My work on this paper has been supported by a Fellowship for Independent Study and Research from the National Endowment for the Humanities (U.S.A.).

REFERENCES

[1] Marilyn Adams, 'Is the Existence of God a "Hard" Fact?', *Philosophical Review*, 76 (October 1967).

[2] John Martin Fischer, 'Freedom and Foreknowledge', *Philosophical Review*, 92 (January 1983).

[3] John Martin Fischer, 'Ockhamism', *Philosophical Review*, 94 (January 1985).

[4] Joshua Hoffman and Gary Rosenkrantz, 'Hard and Soft Facts', *Philosophical Review*, 93 (July 1984).

[5] Nelson Pike, 'Omnipotence and God's Ability to Sin', *American Philosophical Quarterly*, 6 (1964).

[6] Nelson Pike, 'Divine Omniscience and Voluntary Action', *Philosophical Review*, 74 (January 1965).

[7] Nelson Pike, *God and Timelessness* (Routledge, 1970).

[8] Nelson Pike, 'Fischer on Freedom and Foreknowledge', *Philosophical Review*, 93 (October 1984).

[9] William Rowe, *Philosophy of Religion* (Dickenson 1978).

EVIL AND THE EXISTENCE OF A FINITE GOD

By P. J. McGrath

THE problem of evil is almost invariably presented as an objection to the claim that a divine being exists who is both omnipotent and infinitely good. The implication is that to escape the problem one need only alter one's conception of God by limiting his power or his goodness. J. L. Mackie makes this point explicitly:

> It is plain, therefore that this problem can be easily solved if one gives up at least one of the propositions that constitute it. Someone who holds that there is in some sense a god, but one who is not wholly good, or, though powerful, not quite omnipotent, will not be embarrassed by this difficulty. (*The Miracle of Theism*, Oxford 1982, p. 151)

I believe that Mackie is wrong about this and that evil constitutes a problem for belief in even a scaled down version of the deity. To escape the difficulty one would need to reduce the power or the goodness of God to such a degree that he could no longer be properly called God or, at least, could no longer be regarded as a proper object of worship.

Consider first what happens to the problem when God is conceived as infinitely good but not omnipotent. Mackie's thinking is presumably that evil can now be explained by saying that while God does not want evil to be present in the universe, he does not have sufficient power to prevent it. But the trouble with this is that some evils which formerly existed have been eliminated by human ingenuity. For example, smallpox has been wiped out through the use of vaccination. If man can get rid of an evil like smallpox, why could God have not done likewise? To say that he was unable to do so is to reduce his power to such an extent that it is difficult to see how he could be any longer regarded as divine. Certainly a being which lacks the power to do something which man is capable of doing could scarcely be a proper object of worship. The discovery that such a being exists would appear to have no more significance for religious belief than the discovery that there are intelligent beings in some other part of the physical universe.

Our second option is to conceive of God as omnipotent but limited in goodness. This makes it possible to argue that there is evil in the world because God is not as concerned with preventing it as he would have been if his goodness were unlimited. However, as a solution to the problem of evil this is even less satisfactory than the previous move. The trouble now is that a God who tolerated the sort of evils which actually exist simply because of his lack of concern for the welfare of others would be a moral monster. This point can be readily brought out by citing analogies from human behaviour. What would we think, for instance, of a surgeon who performed operations without the use of anaesthetic not because there was no anaesthetic available, but because he was not prepared to take the trouble to see that it was administered? Or how would we look on a doctor who was in a position to cure his patients' ailments, but was so indifferent to their welfare that he refused to take the trouble to do so? An omnipotent deity who tolerated all the evils of the world because of lack of benevolence would be open to much greater moral condemnation than even these two depraved characters – first because the amount of evil involved is so much greater; second because an omnipotent being could eliminate evil by a simple act of will whereas the surgeon and doctor would have had to exert themselves; third because an omnipotent being would be guilty not only of not eliminating the evil which already exists, but, unlike the depraved medics, of not preventing its existence in the first instance. Such a being would appear more akin to an evil demon than to a deity. It would be difficult to see the point, or the propriety, of worshipping him.

There is one final option to be considered. We could abandon both of the propositions which constitute the traditional problem of evil by conceiving God as limited in goodness and power. This would go some distance towards avoiding the difficulties surrounding the second option, since it makes it possible to argue that God tolerates the evils of the world not because he is indifferent to the suffering which they cause, but because he is unable to eliminate them. But as a solution to the problem of evil this is even less acceptable than the first option. For if a being who is infinite in goodness but incapable of eliminating any of the evils of the world is deemed unworthy to be an object of worship, he is clearly even less worthy if he is conceived to be limited in goodness as well as in power.

University College,
Cork, Eire

© P. J. McGrath 1986

On the Plurality of Worlds

DAVID LEWIS

One of America's most influential philosophers here offers a provocative and highly readable defence of modal realism: the thesis that ours is but one of a plurality of worlds and that the individuals that inhabit it are only a few out of all the inhabitants of all worlds. The book will undoubtedly have a profound effect in logic, metaphysics and the philosophies of mind, language and science.
320 pages, hardback **£25.00** (0 631 13993 1)
paperback **£8.95** (0 631 13994 X)

Reasoning with Arbitrary Objects

KIT FINE

This book provides important new insights into logic and the philosophy of language. It offers a semantic account of natural deduction in terms of arbitrary objects, and, in adopting a fresh approach to this, rehabilitates a view that was thought demolished, and gives new methods for the study of generality.
Aristotelian Society Series
228 pages, **£22.50** (0 631 13844 7)

Thoughts

A Theory of Content

CHRISTOPHER PEACOCKE

This book offers a new philosophical theory of thought contents and shows how thoughts can possess truth conditions. Peacocke applies his substantive theory to different types of content, and examines the impact of his theory on current disputes about realism and about the relations between thought and language. It is a philosophical investigation which will have important implications for issues at the centre of modern philosophy of language and mind.
Aristotelian Society Series
*c.*200 pages. *c.***£15.00** (0 631 14674 1)

Needs, Values, Deliberation and Truth

Essays in the Philosophy of Value

DAVID WIGGINS

This collection of essays comprises revised or corrected versions of several published articles, including the author's two well-known studies of Aristotle's theory of practical reason, his attempt to refurbish the incompatibilist view of human freedom 'Towards a Reasonable Libertarianism', and his British Academy Lecture 'Truth, Invention and the Meaning of Life', as well as several previously unpublished papers.
Aristotelian Society Series
(Spring 1986) *c.*240 pages, *c.***£17.50** (0 631 14044 1)

Actions and Events

Perspectives on the Philosophy of Donald Davidson

Edited by ERNEST LEPORE and BRIAN McLAUGHLIN

Donald Davidson is among the most influential philosophers of the twentieth century. This volume includes some thirty essays which variously criticize, comment and develop his philosophy as represented in his collected papers *Essays on Actions and Events*, in addition to three further essays by Davidson himself.
*c.*512 pages, *c.***£29.50** (0 631 14451 X)

Basil Blackwell

108 Cowley Road, Oxford OX4 1JF
Suite 1505, 432 Park Avenue South, New York NY 10016

PHILOSOPHICAL INVESTIGATIONS

Edited by D Z Phillips

Philosophical Investigations features articles, reviews, and in depth critical notices by experts such as Cora Diamond, Frank Ebersole, Antony Flew, Raimond Gaita, Oswald Hanfling, Lars Hertzberg and Rush Rhees.

Volume 9 1986 includes the following articles and reviews:

Articles

Privacy regained
DENNIS M. SENCHUK

The picture theory and assertion
GUY STOCK

Objectivity and meaning: Wittgenstein on following rules
VICTOR J. KREBS

Are we struggling for justice?
SIMONE WEIL

Reviews

Spreading the word, Simon Blackburn
OSWALD HANFLING

Understanding arguments, D. Evans and H. Palmer
MICHAEL COHEN

January, April, July, October
Subscription Rates Volume 9, 1986
Individuals: £14.25 (UK) £17.85 (overseas) US$30.35 (N America)
Institutions: £32.50 (UK) £40.65 (overseas) US$65.55 (N America)

Please send orders with payment to:

Iris Taylor
Journals Department
Basil Blackwell
108 Cowley Road
Oxford OX4 1JF
England

or to: Journals Department
Basil Blackwell
PO Box 1320
Murray Hill Station
NY 10156
USA

Payment can be made by cheque or money order made payable to Basil Blackwell Ltd, by credit card (Access, American Express, Barclaycard, Diners Club, Mastercharge, Visa; please state card number and expiry date), or through the National Girobank (account number 236 6053). Sample copies, single and back issues are available; please write for details.

Basil Blackwell · Oxford and New York

BIOETHICS

a new quarterly journal from Basil Blackwell

CALL FOR PAPERS

In January 1987, Basil Blackwell are launching a new journal, **Bioethics**, which will provide an international forum for the study of ethical issues raised by medicine and the biological sciences. The journal will be interdisciplinary and contributions are equally welcome from those working in philosophy, medicine and the health sciences, law, sociology, economics, public policy, education, and other related fields. Topics in bioethics may be approached from the vantage point of concrete ethical, legal and policy problems, or in terms of the basic ethical concepts, principles and theories used in discussions of such problems.

Information for contributors

1. Manuscripts should be submitted to:

   Helga Kuhse and Peter Singer (Editors)
   Bioethics
   Centre for Human Bioethics
   Monash University
   Clayton
   Victoria
   Australia 3168

   Contributors are recommended to send manuscripts by airmail if possible, as surface mail can take two months.

2. Manuscripts, preferably in triplicate, should be double spaced throughout and typed on one side only on white A4 or American quarto paper. Footnotes should be clearly numbered and follow the text.

3. All papers will be send anonymously to two referees. Contributors should prepare their manuscripts so that after removal of the title page there remains no indication of the identity of the author. It is intended that decisions will be made within two months of receipt of a submitted manuscript.

Basil Blackwell • Oxford and New York

The Philosophical Review

A Quarterly Journal Edited by the Sage School of Philosophy

October 1985

Kant's Theory of Geometry .... *Michael Friedman*

Ways of Moral Learning .... *Richard Miller*

Why Propensities Cannot Be Probabilities .... *Paul Humphreys*

220 Goldwin Smith
Cornell University
Ithaca, New York 14853

U.S. Subscription Rates:	*Rates outside the U.S.:*
$12.00 per year for students, the retired, and the unemployed	$17.00 per year for students, the retired, and the unemployed
$20.00 per year for other individuals; $6.00 per single issue	$25.00 per year for other individuals; $8.00 per single issue
$34.00 per year for institutions; $ 9.00 per single issue	$39.00 per year for institutions; $11.00 per single issue

VISA and MasterCard Accepted

AUSTRALASIAN JOURNAL OF PHILOSOPHY

Contents of Vol. 64, No. 1; March 1985

EDITOR: BRIAN ELLIS

Special Issue on
Universals
and
Volition: Sortal Terms: Induction

ARTICLES	
Donald C. Williams	Universals and Existents
Peter Forrest	Ways Worlds Could Be
David Lewis	Against Structural Universals
John Bacon	Armstrong's Theory of Properties
Stewart Candlish	*Absque Labore Nihil*
E.J. Lowe	Sortal Terms and Absolute Identity
A.A. Derksen	Clendinnen and Salmon on Induction as the Non-Arbitrary Method
DISCUSSION	
D.M. Armstrong	In Defence of Structural Universals
Peter Forrest	Neither Magic Nor Mereology: A Reply to Lewis
David Lewis	Comment on Armstrong and Forrest
John Bigelow	Towards Structural Universals

BOOKS RECEIVED, NOTES AND NEWS, INDEX

ANNUAL SUBSCRIPTIONS: Individuals A$17 (US$18 or Stg£10); Institutions A$34, (US$36 or Stg£18). All correspondence should be addressed to the Editor or Business Manager, Department of Philosophy, La Trobe University, Bundoora, Victoria, 3083, Australia.

PUBLISHED QUARTERLY BY
THE AUSTRALASIAN ASSOCIATION OF PHILOSOPHY

NOTES

The ANALYSIS Committee consists of: Chairman, P. T. Geach; Secretary, T. R. Baldwin; Andrew Harrison, R. F. Holland, J. H. McDowell, D. H. Mellor, M. A. Stewart, R. G. Swinburne, A. R. White, C. J. F. Williams, Peter Winch. This committee is responsible for appointing and advising the Editor and for the general policy of the paper.

SUBSRIPTIONS. ANALYIS is published by Basil Blackwell Ltd, 108 Cowley Road, Oxford OX4 1JF, England and P.O. Box 1320, Murray Hill Station, NY 10156, U.S.A. Each volume comprises four numbers, each of at least 48 pages, appearing in January, March, June and October. Orders should be sent to the publisher. Subscriptions prices for 1986 are as follows:

	U.K.	Overseas	U.S.A.
Institutions	£11.75	£14.50	U.S. $33.00
Individuals	£7.85	£9.75	U.S. $21.00

BACK ISSUES. Volumes 1–37 are now available in complete volumes in an offset reprint from Swetz & Zeitlinger, 471 Keizersgracht, Amsterdam, Holland. Recent back issues are available from the publisher at the current single issue price.

MICROFILM copies of complete volumes are available from Volume 1 onwards from University Microfilms International Ltd, 30–32 Mortimer Street, London W1N 7RA, England, or 300 North Zeeb Road, Ann Arbor, Michigan 48106, U.S.A.

ADVERTISING. For details contact Bill Drake, telephone (0869) 38477, or write c/o the publishers.

U.S. MAILING. Second class postage paid at New York. Postmaster: Send address corrections to: Analysis: c/o Expediters of the Printed Word Ltd, 515 Madison Avenue, New York, NY10022 (U.S. Mailing Agent).

CONTRIBUTIONS. Articles submitted for publication should be addressed to Mr Christopher Kirwan, Exeter College, Oxford OX1 3DP.

Contributors are asked to note the following requirements. Articles should normally not exceed 3,000 words in length, including references and footnotes. They should be type written in double spacing on one side of the paper only; photographic copies are acceptable. Footnotes should be kept to a minimum and wherever possible avoided altogether. Inset quotations and formulae should be well spaced from the text preceding and following them. The spacing of symbols within formulae should be carefully attended to. Single quotation marks should be used, except for purposes of internal quotation and "scare" quotes. Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

Typescripts of rejected articles will be returned only if self-addressed envelopes with correct postage are provided (in the case of overseas contributors, international reply coupons sufficient for airmail).

Proofs of accepted articles will be sent to authors for correction, together with information about free copies. Typescripts of accepted articles will be retained by the Editor on the assumption that authors have kept their own copies.

Apart from correction of printers' errors, only very minor changes in proof will normally be accepted by the Editor, to avoid delay and expense in production.

Typeset by Unicus Graphics Ltd, Horsham. Printed and bound in Great Britain by
Whitstable Litho Ltd, Whitstable

Vol. 46 No. 2 (New Series No. 210) March 198

ANALYSIS

Edited by

CHRISTOPHER KIRWAN

CONTENTS

Quotation and the Liar Paradox — DEAN BUCKNER *and* PETER SMITH
On set theoretic possible worlds — CHRISTOPHER MENZEL
An observation on common names and proper names — JOHN TIENSON
Prior's analytic — B. H. SLATER
'If A and B then A' — A. J. DALE
Reply to Dale — E. J. LOWE
Transparently false: reply to Hardin — D. K. BUCKNER
AE or EA? — H. W. NOONAN
Hume's argument for the temporal priority of cause over effect — MICHAEL J. COSTA
The immateriality of 'abstract objects' and the mental — IRVING THALBERG
Churchland on direct introspection of brain states — NATIKA NEWTON
Nozick, justice, and the sorites — ROY A. SORENSEN
Harris on non-vicious choice — OLIVER LEAMAN
Pascal's wager and infinite utilities — ANTONY DUFF
Repeatable miracles? — ANDREW REIN
Correction

BASIL BLACKWELL PUBLISHER · OXFORD

ISSN 0003-2638

Actions and Events

Perspectives on the Philosophy of Donald Davidson
Edited by ERNEST LEPORE and BRIAN McLAUGHLIN

Donald Davidson is among the most influential philosophers of the twentieth century. This volume includes some thirty essays which variously criticize, comment and develop his philosophy as represented in his collected papers *Essays on Actions and Events*, in addition to three further essays by Davidson himself.
528 pages, **£29.50** (0 631 14451 X)

Truth and Interpretation

Essays on the Philosophy of Donald Davidson
Edited by ERNEST LEPORE

Truth and Interpretation comprises twenty-seven critical essays, a substantial introduction to Donald Davidson's philosophy of language, and three essays by Davidson himself. The volume's five sections correspond to the major sections of his *Inquiries into Truth and Interpretation*. Each contains critical essays addressing, interpreting and further developing his views
536 pages, **£29.50** (0 631 14811 6)

Thoughts

A Theory of Content
CHRISTOPHER PEACOCKE

Christopher Peacocke offers a new philosophical theory of thought contents and shows how thoughts can possess truth conditions. He applies his theory to different types of content, and examines it's impact on current disputes about realism and the relation between thought and language
192 pages. **£15.00** (0 631 14674 1)

On the Plurality of Worlds

DAVID LEWIS

One of America's most influential philosophers here offers a provocative and highly readable defence of modal realism: the thesis that ours is but one of a plurality of worlds and that its inhabitants are only a few out of all the inhabitants of all worlds.
288 pages, hardback **£25.00** (0 631 13993 1)
paperback **£8.95** (0 631 13994 X)

Wittgenstein: Rules, Grammar and Necessity

An Analytical Commentary on the *Philosophical Investigations*, Volume 2
G. P. BAKER and P. M. S. HACKER

Consisting of exegesis and philosophical essays, this long-awaited second volume covers *Investigations* 185–242, the highly controversial remarks on following rules which lead up to the private language argument. Baker and Hacker demonstrate that the most commonly held interpretations of this aspect of Wittgenstein's philosophy are deeply misguided.
368 pages, **£29.50** (0 631 13024 1)

Basil Blackwell

108 Cowley Road, Oxford OX4 1JF
Suite 1505, 432 Park Avenue South, New York NY 10016

The Monist

Founded 1888 by **Edward C. Hegeler**

An International Quarterly Journal of General Philosophical Inquiry

Editor, John HOSPERS

Editorial Board: William P. Alston, Monroe C. Beardsley, Dagfinn Føllesdal, R. Barcan Marcus, Joseph Margolis, Richard Martin, Wallace I. Matson, Joseph Owens, Richard Rorty, Wilfrid Sellars, Anita Silvers, John E. Smith, Richard Wasserstrom.

Managing Editor: Sherwood J. B. SUGDEN

EACH ISSUE IS LIMITED TO ARTICLES ON A SINGLE GENERAL TOPIC. Communicate in advance with the Editor for special instructions. Two nonreturnable copies (one copy outside the U.S. and Canada) should be submitted NINE MONTHS PRIOR to the announced publication date. Papers should be prepared for anonymous submission to readers, and should be from 10 to 20 double-spaced typewritten pages in length, including notes.

SCHEDULED PUBLICATION DATES:

Vol. 68, No. 2	Apr.,	1985	Knowledge, Justification, and Reliability (Part II)
Vol. 68, No. 3	July,	1985	The Concept of Law
Vol. 68, No. 4	Oct.,	1985	The Nature of Experience
Vol. 69, No. 1	Jan.,	1986	Philosophy and Literary Theory
Vol. 69, No. 2	Apr.,	1986	Logic and Philosophy
Vol. 69, No. 3	July,	1986	The Nature of the Soul
Vol. 69, No. 4	Oct.,	1986	Intentionality
Vol. 70, No. 1	Jan.,	1987	Animal Rights

Editorial Office: School of Philosophy, University of Southern California
Los Angeles, CA 90089
Business Office: Box 600, La Salle, Illinois 61301

SUBSCRIPTION RATES: United States: Annual (4 issues): Institutions, $24.00; Individuals, $15.00. Single copies: $6.00. Foreign postage: Add 50 cents to single copy rate or $1.00 to subscription.

SOCIAL PHILOSOPHY AND POLICY

Edited by Ellen Frankel Paul

Each 192-page issue of **Social Philosophy and Policy** is devoted to a single topic in social or political philosophy, with specially invited papers by leading scholars on the most important issues of our time. The topics for Volume 3 are **Nuclear Rights/Nuclear Wrongs** and **Marxism and Liberalism.**

October and April for the Social Philosophy and Policy Center, Bowling Green State University, USA

Subscription to Volume 3, 1985–6
Individuals: £10.50 (UK) £12.00 (overseas) US$14.50 (N America)
Institutions: £24.50 (UK) £29.50 (overseas) US$40.00 (N America)

Please send orders with payment to:

Iris Taylor
Journals Department
Basil Blackwell
108 Cowley Road
Oxford OX4 1JF
England

or to: Journals Department
Basil Blackwell
PO Box 1320
Murray Hill Station
NY 10156
USA

Basil Blackwell • Oxford and New York

QUOTATION AND THE LIAR PARADOX

By DEAN BUCKNER *and* PETER SMITH

LAURENCE Goldstein [5] purports to give a solution of the Liar Paradox by showing that the identity statement

(1) *k* = ‘*k* is not true’,

which generates the Paradox, is simply false.

At first sight, this ‘solution’ merely gives rise to another paradox. For how can (1) possibly be false? – it is surely intended to record a *stipulation* and is thus true by definition. The expression on the right is a designator whose reference has been fixed in advance, and the expression on the left is an as-yet-uninterpreted symbol, whose reference is to be fixed via the identity. It seems, therefore, that (1) is analogous to a statement like

(2) j = 13,

which one could use in mathematics to stipulate the reference of the constant ‘j’, or like

(3) A = Aristotle,

where the symbol ‘A’ has been made arbitrarily to refer to a well known philosopher of antiquity. But neither (2) nor (3) is obviously illegitimate. What makes (1) any different?

Goldstein allows that we can assimilate (1) to (2) and (3) if we can construe quotation-expressions on strict Tarskian lines, i.e. as semantically unstructured names. On the pure Tarskian view, the ‘*k*’ within the quotation-expression “‘*k* is not true’” functions just like the ‘A’ in ‘Aristotle’: it is then of no significance that the symbol ‘*k*’ occurs on both sides of (1) any more than it matters that the letter ‘A’ occurs on both sides of (3). And although Goldstein doesn’t mention the point, he would presumably agree that the same holds if quotation is construed as abbreviated structural description. On this familiar neo-Tarskian account, the quotation expression in (1) is first unpacked as

The result of concatenating ‘*k*’, ‘ ’, ‘i’, ‘s’ . . . ,

and the residual quotation-expressions are then taken as unstructured names in the manner of the pure Tarskian account. If the quotation in (1) is treated in this manner, then again the fact that '*k*' occurs on both sides creates no problems.

Goldstein maintains, however, that the strict Tarskian theory of quotation – and no doubt the structural description theory as well – is wrong: the correct theory is in fact a Davidsonian or *demonstrative* one (see [3], [4]). On this theory, he argues, it is evident that (1) is *not* legitimate.

> Consider a person, X, who makes the statement '*k* is not true'. Here '*k*' names a statement said by X to be not true. So X is asserting a statement that predicates non-truth of *k*. Hence it is not for X simply to 'let' '*k*' be the name of *that* statement, for he has *already* assigned the name '*k*' to a statement he has said is *not* true, namely *k*. ... The restriction on designation-assignment that we have discovered arises only out of a simple consistency requirement: if '*a*' designates a statement and '*b*' designates a negation of that statement, then $a \neq b$. ([5] p. 11)

It is easy for a theoretician to overlook this restriction, he adds, since the stipulation in (1) *would* appear legitimate from the perspective of a classical Tarskian theory of quotation.

> However, the Tarskian theory of quotation is false. Quotation marks do not serve to form semantically unstructured names, but are *pointing* devices, and in the quotation-expression '*k* is not true' [*sic*] as we have been using it, '*k*' is pointed to as the designator of a statement of which non-truth is predicated. Hence the restriction on using '*k*' for some incompatible purpose. (p. 12).

In short, then, the Davidsonian theory of quotation makes the stipulation (1) illegitimate, and the argument to the Liar Paradox cannot even get started.

We have two problems about all this. First, suppose we grant Goldstein his two theses that the quotes in (1) function according to the Davidsonian theory, and that *because* of this (1) is simply false. What prevents us introducing another quotational device, symbolized by e.g. a pair of asterisks, which functions in a Tarskian way – or better, if we are to avoid an infinite lexicon, in the neo-Tarskian way? In other words, why shouldn't we propose that '**k* is not true*' is to function as a structural description of the sequence of symbols: *k* is not true. There can be no objection, Davidsonian or otherwise, to this procedure. But in that case

(1t) $k = {*k}$ is not true*

legitimately stipulates that '*k*' also denotes that sequence, and the Liar Paradox has been reinstated in a version that, by hypothesis, is not amenable to Goldstein's treatment. Anyone who now utters '*k* is true' has (apparently) said that *k* is not true – which gives us the Paradox in one of its traditional forms.

Thus Goldstein's theses establish, at the very most, that a Davidsonian reading of the quotation marks in (1) renders the Liar generator illegitimate. But any serious treatment of the Liar Paradox must surely cover a much wider family of cases, and deal with Liar generators constructed with other devices (structural descriptions, Gödel numberings, or whatever). Our first complaint, therefore, is that Goldstein's approach cannot advance our *general* understanding of the Liar family.

Second, it is quite unclear that Goldstein is right even about the implications of a Davidsonian theory of quotation. Adapting his 1984 account of the theory (see [4]), we are to read (1) as

(1d) $k =$ the type of which a token is displayed here: k is not true.

Exactly what is supposed to be illegitimate about *that*?

Suppose you start off by considering the sentence type '*k* is not true', with as-yet-undefined '*k*'; you then (attempt to) fix the reference of '*k*' by offering (1d). And finally you go on to wonder if *k*, so defined, is actually true, and find yourself entangled with the Liar. Is there any stage at which, like the X of Goldstein's argument, you *assert* the statement '*k* is not true'? Not at the initial stage, surely, when you are considering the not-fully-interpreted *sentence* '*k* is not true'. Nor at the final stage when, wondering about the truth-value of *k*, you proceed suppositionally: e.g. 'If *k* is true, then ... '. In between, you *do* assert (1d) – but of course it is not part of the Davidsonian theory that this involves asserting the ostensively displayed contents. To emphasise a point made by Buckner ([1], [2]), the asserted part of (1d) is '$k =$ the type of which a token is displayed here'. The expression following the colon in (1d) is not to be understood as part of the statement made, any more than the fish you hold up while uttering 'I caught this fish today' is a part of what you are actually *saying*. Hence at no stage in the imagined procedure do you in fact assert '*k* is not true'. So you do *not* (assertorically) predicate non-truth of *k*, and may therefore generate the Liar without committing the alleged offence against Goldstein's consistency requirement.

Finally, if you read the symbols flanking the identity sign in (1) as specifying statement contents, i.e. if you read (1) as

(1s) What k states $=$ what 'k is not true' states.

then this surely *is* false, in virtue of the meaning of the truth-predicate. Perhaps this is Goldstein's intended point. But what has this to do with a *Davidsonian* theory of quotation? And even if we allow (1s) as the preferred reading of (1), there will still remain Liar generators like (1t) to deal with. Hence, so far as Goldstein's arguments are concerned, the Liar remains very much in business.

© DEAN BUCKNER *and* PETER SMITH 1986

University of Bristol
9 Woodland Way
Bristol BS8 1TB

and

University College of Wales,
Aberystwyth,
Dyfed SY23 3DY

REFERENCES

[1] D. K. Buckner, 'Goldstein on Quotation', *Analysis*, 44.4, October 1984, pp. 189–90.
[2] D. K. Buckner, 'Reply to Van Brakel', *Analysis*, 45.4, October 1985, pp. 215–17.
[3] D. Davidson 'Quotation', *Theory and Decision*, 11 (1979), pp. 27–40, reprinted in his *Inquiries into Truth and Interpretation*, Oxford: Clarendon Press, 1984.
[4] L. Goldstein, 'Quotation of Types and Other Types of Quotation', *Analysis*, 44.1, January 1984, pp. 1-6.
[5] L. Goldstein, 'The Paradox of the Liar – A Case of Mistaken Identity' *Analysis*, 45.1, January 1985, pp. 9–13.

ON SET THEORETIC POSSIBLE WORLDS

By CHRISTOPHER MENZEL

IN a recent note [3], Selmer Bringsjord has argued that the set theoretic definition of possible worlds proffered by, among others, Robert Adams and Alvin Plantinga is incoherent. It is the purpose of this note to evaluate that argument.

According to Adams, the notion of a possible world can be analysed into that of a *world-story*, i.e., 'a set which has as its members one of every pair of mutually contradictory propositions, and which is such that it is possible that all of its members be true together' ([1], p. 204). Plantinga doesn't actually commit himself to the identification of possible worlds with such sets, but he does agree that there is such a set corresponding to each possible world W, which he calls the *book on W* ([5], pp. 44–6). Now, according to Bringsjord, the trouble with world-stories is this. Let S be any world story. S has some cardinal number k. Consider then the power set P(S) of S. By Cantor's theorem, card (P(S)) = $2^k > k$. For each member s of P(S) there is the proposition $p(s)$ that s is a set. (It is assumed here, and henceforth, that the mapping p is one-to-one.) Since each such proposition would have been true had S been actual, it follows that for every $s \in P(S)$, $p(s) \in S$. Thus there is a one-to-one mapping that takes each element of P(S) to an element of S. But then by an elementary theorem of set theory, card (P(S)) = $2^k \leqslant k$ = card(S), contradiction.

The argument is conclusive. Clearly, if the notion of a world-story can be salvaged it must undergo revision. One possibility suggests itself immediately: world stories, though collections of *some* kind, are not sets. Indeed, the world-story theorist might reply, this is less a revision than a correction, for it is obvious world-stories *can't* be sets once one remembers that the necessary propositions of mathematics are among the true propositions there are. In particular there is, for each set *m* in the collection V of pure sets, the proposition that, e.g., *m* is nonselfmembered. Since each such proposition is necessarily true, each is a member of every world-story, and in particular of S. It follows that V can be injected into S. But, as Bringsjord well notes, if S were a set it would have a cardinal number k, from which it would follow that V has a cardinal number $\leqslant$ k. That, however, is provably false. So there are world-stories all right, it's just that they are 'too big' to be sets. Thus, Bringsjord's argument fails at a crucial juncture, viz. in its initial assumption that the world story S is a set and hence has a definite cardinality.

This is too quick, though. For the assumption that S has a definite cardinality is not essential to Bringsjord's argument. What *is* essential is the claim that, whatever sort of collection S may be, there is a collection of all the subcollections of S. From that assumption, one can deduce the denial of Cantor's theorem, which in its most common form says nothing about cardinality (it says only that S can be mapped one-to-one into P(S), but not vice versa) and applies generally to any collection that has a corresponding 'power collection'. Thus we needn't suppose that S has a cardinal number for Bringsjord's argument to go through; we only need to suppose that it has a power collection.

But isn't this assumption entirely unjustified? Don't we know that collections like S that are 'too big' to be sets don't have corresponding power collections? I don't think the world-story theorist can get off the hook so easily. On what, exactly, does he base his claim that 'large' collections don't have power collections? Granted, in class theories like Von Neumann-Gödel-Bernays (VNBG) in which large collections like V are allowed, V and its ilk do not have power collections. But we can't conclude from this fact alone that there is no such thing as P(V), or more generally P(C) for any large collection C, any more than we can conclude there is no such thing as V from the fact that V doesn't exist with respect to Zermelo-Fraenkel set theory (ZF). At the least, if the world-story theorist is going to marshal results in formal set theory to do the metaphysical labour they are intended to do here, more argument is needed. Now, for some large collections at least, such argument is available; it will not, however, be of help to the world-story theorist. Let me elaborate.

The central reason for denying that 'proper classes' like V have corresponding power collections is not simply that they are too big to be sets, but rather that there is a crucial *structural* difference between sets and proper classes. The conceptual foundation of contemporary set theory is the so-called *iterative* conception of set. On this conception, intuitively, a set is any collection that can be 'built up' from an initial collection of urelements by the repeated application of certain set-building operations, e.g., in ZF, pairing, power set, union, replacement, and separation. Given some particular choice of urelements, the resulting closure under the chosen operations results in a cumulative hierarchy of levels, one for each ordinal number, each of which contains all the collections that can be formed from the elements of all the previous levels. What is distinctive about V on this conception is not that it is bigger than any set can be in ZF or VNBG, but rather that it cannot itself be built up in the way a set can, since it is precisely the *ultimate result* of such a building-up process (beginning, of course, with an empty initial collection of urelements). In more formal terms, V is *unbounded*, that is, it contains members of arbitrarily high rank (the rank of an object is defined recursively to be the least ordinal greater than the ranks of all its members); i.e., one might say, of arbitrarily high complexity with respect to the set-building operations. It is this feature, and not size, which is the true conceptual boundary between sets and proper classes. (I have argued this at length in [4].) Thus one might argue that the reason V, and proper classes generally, have no power collection is that they lie 'beyond' the range of any set building operation; they are of an entirely different nature from sets altogether. Note, though, that on this characterization of the distinction between sets and proper classes, which has nothing to do with *size*, there is no basis for denying that a given collection is a set just because it is larger than any set can be in ZF. In particular, on this way of slicing things up, our world-story S is indeed a set after all (though of course not in the sense of ZF): since its elements are all propositions, and propositions for the world-story theorist are not sets, they themselves have no members and so have a rank of 0; hence, S itself has a rank of 1, i.e. S is bounded. Thus there is no more reason to deny that S has a power set than there is to deny that any set does.[1] But then, of course, our revised version of Bringsjord's argument appears inescapable.

Is it possible to circumvent the argument in some other way? There seem to be only two alternatives, neither (at first blush, anyway) very appealing. The first is to deny the existence of the propositions $p(s)$. This, however, it tantamount to abandoning the notion of a world-story. To all appearances, the propositions in question are simple instances of the following unimpeachable principle: if A is any property and x is any object, then there is the proposition that x has A. It is difficult to imagine there being a nonarbitrary way of ruling out the $p(s)$ as genuine instances of the principle that wouldn't simultaneously rule out virtually any other proposition of the same form. And without such 'atomic' propositions, it is hard to see how there could be any propositions at all.

---

[1] Of course, large sets cannot have a definite *cardinality*; rather, they are, in Cantor's terms, mathematically indeterminable. The burden of [4] is to show that unboundedness and mathematical indeterminability are entirely distinct properties. To demonstrate this I have developed a ZF-style set theory in that paper which admits sets that are too big to be sets in the sense of ZF and VNBG (in particular, there is a set of all ordinals, where these are *not* construed set theoretically à la von Neumann) and have proved its consistency relative to ZFC with an inaccessible cardinal.

The second option is to deny the full power set axiom. The problem with this move, though somewhat more palatable than the first, is that it seems wholly out of character with the strong realism that motivates the world-story theorist's project in the first place. Admittedly, those of a more constructivist turn have reason to be sceptical about so powerful an axiom. However, one would think that for anyone so realistic as to admit an ontology of sets and propositions sufficiently rich to attempt the sort of reconstruction at issue, it would be awkward indeed to have to deny a proposition which is a hallmark of realism in the philosophy of mathematics.

But would it? Though I tend to agree with the reasoning in the previous paragraph, I'm not so sure that it does justice to the proposed option. The assumption there seemed to be that a strong logical realism about properties, relations, and propositions ought to carry with it an equally strong realism about the existence of sets, in particular the sort of realism that buys into the powerful, highly nonconstructive set-building operations of ZF, especially the power set axiom. But one might well argue that logical realism has no clear implications at all regarding set theoretic realism (other than perhaps a commitment to *some* sort of realism or other, however weak), and that in fact a more conservative set theory like ZF-Power (or even a predicative theory like Kripke-Platek set theory with urelements (KPU); v. e.g. [2]), which maintains a closer connection between expressible properties and the existence of sets, is a more natural mathematical companion to logical realism than full ZF. By adopting ZF-Power (or some similarly conservative set theory), then, and adjusting it appropriately to allow for the existence of large sets[2] (and urelements), the world-story theorist is free to postulate the existence of his world-stories without fear of paradox, at least by way of Bringsjord's argument; for that argument depends essentially on there being a power set of the world-story S; but there is simply no way of generating the full classical power set of an infinite set in ZF-Power (let alone KPU); the most one can hope for is a set of subsets that is equinumerous to the original set, and that clearly won't do the job.

So the upshot of Bringsjord's argument is clear: the world-story theorist can have either his world-stories or an iterative understanding of sets that includes the power set axiom, but not both. For on the iterative conception, there is no reason to deny that world-stories are sets, despite their large size. But then, given the power set axiom, Bringsjord's argument is unavoidable.

---

[2] This could be done essentially as in [4], viz. the large sets are defined to be just those sets which are not in one-to-one correspondence with any pure set, i.e. any set whose transitive closure contains only sets. Restriction of the axiom schema of replacement to 'small' sets is also necessary.

A caveat. That world-stories, the iterative conception, and the power set axiom together create a volatile metaphysical mixture doesn't mean that *every* conception of possible worlds along these lines is in a similar conceptual fix. It is not obvious, for example, that Plantinga's construction in terms of states-of-affairs falls prey to similar objections. Also, perhaps the notion of a world-story itself could be amended by appropriately 'filtering out' necessary truths and complex truth-functional contingent propositions that are entailed by more 'basic' propositions (and hence which play no essential role in delimiting any particular 'way things could have been') and constructing world-stories from the (countably many?) propositions that remain. A world-story would than be a set S that is maximal in the sense that for any proposition q, S *entails* (but not necessarily *contains*) either q or its negation, but not both (where of course S entails q just in case it is not possible that all the members of S be true without q being true). This is certainly an avenue worth exploring. Finally, I should mention a metaphysical theory put forward by Edward Zalta in [6] that offers both a rigorous reconstruction of possible worlds and a realistic but non-standard understanding of what sets are that enables one to embrace both the iterative conception and power set while skirting the difficulties at issue here.[3]

© CHRISTOPHER MENZEL 1986

Center for the Study of Language and Information,
Ventura Hall, Stanford University,
Stanford, California 94305, U.S.A.

REFERENCES

[1] Adams, Robert M., 'Theories of Actuality,' in Michael J. Loux (ed), *The Possible and the Actual*, Cornell 1979, pp. 190–200.
[2] Barwise, Jon, *Admissible Sets and Structures*, Springer 1975.
[3] Bringsjord, Selmer, 'Are there Set Theoretic Possible Worlds?', *Analysis*, 45.1, January 1985.
[4] Menzel, Christopher, 'On the Iterative Explanation of the Paradoxes,' forthcoming in *Philosophical Studies*.
[5] Plantinga, Alvin, *The Nature of Necessity*, Oxford 1974.
[6] Zalta, Edward, *Abstract Objects: An Introduction to Axiomatic Metaphysics*, Reidel, 1983.

---

[3] I would like to thank an anonymous referee for helpful comments on an earlier draft of this paper.

AN OBSERVATION ON COMMON NAMES AND PROPER NAMES

By JOHN TIENSON

> Common names, for Mill, have both connotation and denotation. Thus 'horse' connotes certain properties, and the name 'horse' denotes the things that have those properties. By contrast, proper names have no connotations; they do not denote in virtue of the possession of certain properties by their denotations, but so to speak, directly. Thus Socrates received his name by being dubbed 'Socrates'; and he might just as well have been given any other name.
>
> This contrast is misleading. After all, we might have named horses by another name, too; e.g., 'cow' or 'Pferd'. However, once the convention by which they are called 'horses' is established, it is not correct to call them 'cows'. A horse is not a cow. Just so Socrates could have been named 'Plato' or 'Moses', but once he has been named 'Socrates', it is just as wrong to call him 'Plato' as it is to call a horse a 'cow'. What is correctly called a 'horse' is so called in virtue of its possession of certain properties, just as what is called 'Socrates' is so called in virtue of his possession of the requisite properties. From this point of view, proper names are words like any others. (Leonard Linsky, *Oblique Contexts*, University of Chicago Press 1983, pp. 16f.)

THE view that proper names, like general terms, apply to things on the basis of their properties is a common one. But there is a fundamental difference between general terms and proper names. Because of this difference proper names are not 'words like any others', and Mill's contrast is not misleading. But it is – at best – misleading to say 'what is called "Socrates" is so called in virtue of his possession of the requisite properties.'

General terms apply to new instances. If you know the meaning of a general term, normally you can apply it to things you encounter for the first time. If you could not tell whether some new object was a horse, or brown, that would be good evidence that you did not understand the word 'horse' or 'brown'.

What is true of individual speakers is true for all speakers collectively. There are mountains and brown things on Mars – that is, things of which the words 'brown' and 'mountain' are true. No one has been to Mars to call them 'brown' or 'mountain'. This is an obvious, frequently noted fact about general terms. It follows that general terms are not literally common names. Things are not given the general terms that apply to them in the way they are given their proper names, for general terms apply to countless things to which they could not have been given.

For a general term to apply to something that has never been encountered by any speaker, it must apply on the basis of some relationship between that thing and things to which the term has already been applied, what it has in common with them or its similarities to them. There must be something that makes a Martian mountain a mountain. What makes it a mountain cannot be that it was named 'mountain', because it wasn't. What makes the word 'mountain' true of it must be something that it shares with other mountains, that is, some of its properties or relations. Because general terms apply to new instances, there is no denying that they apply in virtue of properties.

Proper names, in contrast, are given to specific things. They do not apply to things other than those to which they have been given as names. Thus, we do not have the same reason to say proper names apply on the basis of properties as we have in the case of general terms. The reason is absent in the case of proper names. The cases are not in the same boat, as Linsky implies.

Do proper names apply on the basis of properties? Merely noting this difference between proper names and general terms certainly does not prove that they don't. But I think we can see that this is a difference that makes a difference. Why is it correct to call this pencil 'yellow'? Because it has a certain property, not because it was given the name 'yellow'. And why is it correct to call this animal 'horse'? Again, because it has certain properties, although it is more difficult to say just what those properties are. But why is it correct to call this child 'Heather' and that one 'Renee'? Because those are the names they were given!

It would be wrong to say that we call this animal 'a horse' because it was given the common name 'horse'. It would be just as wrong to say we call this person 'Heather' because of her long blond hair and that one 'Renee' because of her short dark hair, or anything of the sort, no matter how we expand the list of properties. It would be wrong, more generally, to say we call her 'Heather' because she has the properties by which we recognize her as Heather. We call her 'Heather' because that is her name, and it is her name because it is a name she was given.

But sometimes we call someone something because of their properties. In the sense in which this is true, it illustrates the point, rather than counting against it. For it means we *give* them names because of their properties. Why is it correct to call this man 'Curly', and not correct to call that, equally bald, man 'Curly'? Because the first has been given the nickname 'Curly', and the second has not. He has been nicknamed 'Curly' because he is bald, and he has this name because it was given to him. There is no reason to conflate the two. He would not have the name if it had not been given to him, no matter how much or little hair he had. And he might retain the name even if he lost the property that led us to give it to him in the first place.

Perhaps we would have no use for a name if we could not recognize its bearer. And no doubt we do recognize things by their properties and relations. But that does not mean that the name applies in virtue of those properties and relations. The name applies to a certain thing because it has been given to that thing. We recognize the thing by its properties. There is no reason to conflate these two things, either.

Compare. Perhaps we would have no use for a name if we had no interest in its bearer. And no doubt there are some things that we are interested in only because they are useful. We should not conclude that the names of such things apply in virtue of the uses of their bearers.

It may also be the case that a thing is the particular thing it is by virtue of its possession of certain properties. Without those properties, it would not be that thing. If some such essentialism is true, it does not follow that names apply to things in virtue of their essential properties. We can, and I think should, still say that a name applies solely because it has been given to a thing, not in virtue of the properties which make it that thing.

Once the relevant conventions have been established, Linsky says, it is wrong to call Socrates 'Plato', and it is likewise wrong to call horses 'cows'. *This* is misleading. 'The convention by which they are called "horses" is a convention of calling all things of that kind 'horses'. The convention by which she is called 'Heather' is a convention of calling this particular person 'Heather'. It is not a convention of calling the sole member of some mighty specific kind by that name. Her properties let us recognize her to call her by name, but that does not alter the nature of the convention, according to which the name attaches to the one to whom it was given, not to properties.

The role of properties in our thought about individuals, important though it is, is not in the semantics of proper names. It is not possible to maintain that properties play no role in the semantics of general terms. This difference between general terms and proper names makes Mill's doctrine that general terms have connotations while proper names are direct connotationless designators a reasonable one.

However, many philosophers today hold that certain general terms, in particular, natural kind terms, are direct designators. This view is sometimes explained by saying natural kind terms lack connotation (e.g. Nathan U. Salmon, *Reference and Essence*, Princeton University Press 1981, Introduction, Chapter 2), and is clearly meant to imply that natural kind terms are importantly analogous to proper names as Mill understood them. What are we to make of this? We cannot, with any hope of success, take it to mean that natural kind terms do not apply to things in virtue of their properties. What is meant, of course, is that natural kind terms do not apply in virtue of properties *that speakers associate with the word*. A natural kind term like 'water' applies to all things of a certain kind. Being of the same kind is determined by the nature of things. And it is an empirical fact that things have hidden and minute properties or natures which we can take to determine their kinds. As Putnam said, 'The key point is that the relation [of being the same kind of liquid] is a theoretical relation: whether something is or is not the same liquid as this may take an indeterminable amount of scientific investigation to determine' ('The meaning of "meaning"', in Hilary Putnam, *Philosophical Papers II: Mind, Language and Reality*, Cambridge University Press 1975, p. 225).

Properties determine sameness of kind, and hence the application of natural kind terms like 'water'. But what properties do this may not be known to a user of the word, or to any speaker. Natural kind terms are used with a faith that nature will provide properties that can be regarded as determining the kind of question. It seems that nature does not often let us down. Thus, the fact that general terms apply on the basis of properties does not imply that they apply on the basis of properties speakers associate with them.

I have tried to make two points in this paper. The main point is, since general terms apply to new instances and proper names apply only to things to which they have been given, it is necessary to say that general terms apply on the basis of properties, but reasonable to say proper names do not. Second, it does not follow that general terms apply on the basis of properties speakers associate with them.

Plainfield, New Jersey, U.S.A. ©JOHN TIENSON 1986

PRIOR'S ANALYTIC

By B. H. SLATER

IN his posthumously published book *Objects of Thought* (Oxford 1971) Arthur Prior develops out of ordinary propositional and predicate logic (see p. 101) a logic of expressions like '. . . fears that . . . ' and ' . . . thinks that . . . ' which 'are as it were predicates at one end and connectives at the other' (p. 19). This construal of propositional attitudes enables Prior to avoid the reification of propositions which is so objectionable in intensional logic (see ch. 3), but also to avoid the abandonment of propositions which would be necessary in a completely extensional logic. Prior's approach even generates some theorems about ' . . . means that . . . ' (p. 105) and so is ambitious enough to raise the hope of constructing a language entirely containing its own semantics. However, the book is incomplete as it stands, being compiled, in part, from manuscript notes, and on the latter possibility, in a chapter the editors (p. viii) describe as an 'afterthought', Prior ends by saying 'on the whole the prospects of a language containing its own semantics still don't look too bright'. I think Prior was too pessimistic in this judgement: there is a slip in his argument, remedying which considerably alters the picture.

On pages 85-86 Prior says,

> ... it can be rigorously proved, using only propositional calculus and quantification theory, that if X says in his book that something that he says in his book is false, then something that he says in his book *is* false, i.e. his statement is bound to be true ... We cannot assert in a book that something asserted in the book is not the case unless something other than this both is asserted in the book and is not the case. No doubt in a book in which nothing else is asserted but truths, or indeed in which nothing else is asserted at all, anyone can inscribe, or have inscribed the *sentence* 'Something asserted in this book is false', but he cannot then *say* by his inscription, or by any inscription, what would normally be said by this one, namely *that* something asserted in the book is not the case. He, or anybody else, could say this very thing elsewhere, and would then, of course, be saying it falsely, but he *cannot* (under these conditions) say it there.

This latter – not being able to say it there – is a bizarre possibility and more disastrous to Prior's explicit theory than Prior allows. After its re-statement later (p. 103) he muses (p. 104) 'Roughly, I think, we use a sentence to say that p if we use it with some degree of isolation and the sentence means that p ... At the same time we can also get into trouble in some cases through *too much* isolation, as in the paradoxical cases we have been mulling over. Or perhaps in these cases the sentence doesn't *mean* what it ordinarily means'. But this misses the crucial point. If Prior cannot provide a *syntactic* characterization of the sentences his 'propositional quantification' is to range over, if, for instance, 'Something X says in X's book is false' does not express a proposition if it is alone in X's book, then Prior's whole analytic falls to the ground. For then the relevant theorem (see p. 103)

$$Sx(\exists p)(Sxp \;\&\; {\sim}p) \rightarrow (\exists p)(Sxp \;\&\; {\sim}p)$$

is not formally available by substitution from Prior's general theorem 'A' (p. 86, where 'δ' should be inserted after the first '*K*'). Indeed, formal substitution is not available, period (see p. 88).

Fortunately for Prior, however, formal substitution *is* available in his system, and Prior's lapse lies elsewhere: in the belief that X cannot say by his lone inscription what would normally be said by that inscription. The point hinges on separating 'asserting 'p'' from 'asserting that p' – a distinction which Prior himself is at pains to make (see ch. 2 and the above quote). The rule about 'assertions' is 'We cannot assert in a book that something asserted in the book is not the case unless something other than this both is asserted and is not the case'. If the 'things' asserted are taken to be *sentences*, this looks as though it implies there must be at least two *sentences* in the book, for 'Something asserted in this book is false' to mean what it normally means, and this is the way Prior, despite his general strictures, seems to have taken it. What the rule really means, however, is that what is involved in saying in a book that something asserted in that book is false is not that 'p' and 'q' are there asserted, where 'p' ≠ 'q', but that there it is asserted that p and asserted that q, where p ≠ q, i.e. that two *propositions* are expressed the book – and this, of course, can be achieved, in an extreme case, by one sentence, so long as that sentence is *ambiguous.*

The possibility of ambiguity Prior only comes to at the end of his discussion, in his 'afterthought'. Perhaps he is not persuaded of its importance because he still hankers after a Tarskian language (p. 104). He does make one pertinent comment, though, (p. 106) 'We could then say that if x means that x is false it will have two contradictory meanings – that it is false and that it is true', but his repetition of 'x' here means that he has to spoil this insight immediately, by adding 'and will be simply and non-paradoxically false' – he forgets that in his non-Tarskian approved grammar sentences cannot be false (or true) and that, on the given supposition, *two propositions* are involved, one false the other true. The properly expressed theorem Prior enunciates on page 105: if x means that something that x means is false there are at least two distinct things that x means. Since this matter has been so controversial I will illustrate this theorem with two formal proofs.

I use 'Msp' for 's means that p' and consider first s_1:

This sentence expresses a false proposition.

With regard to this we can say (truth being now a modality)

$$Ms_1\,(\exists p)(Ms_1p \,\&\, {\sim}\,Tp)$$

and ask whether $T(\exists p)\,(Ms_1p \,\&\, {\sim}\,Tp)$ or ${\sim}\,T(\exists p)\,(Ms_1p \,\&\, {\sim}\,Tp)$. Taking the latter alternative first, and using the new 'T-schema' '$(p)\,(Tp = p)$' (see p. 54 for the laws of propositional identity) we get

$${\sim}(\exists p)\,(Ms_1p \,\&\, {\sim}p)$$

i.e. $(p)\,(Ms_1p \rightarrow p)$

giving $Ms_1\,(\exists p)\,(Ms_1p \,\&\, {\sim}\,Tp) \rightarrow (\exists p)\,(Ms_1p \,\&\, {\sim}\,Tp)$

so $(\exists p)\,(Ms_1p \,\&\, {\sim}\,Tp)$

i.e. $(\exists p)\,(Ms_1p \,\&\, {\sim}\,p)$

and so CONTRADICTION

Hence that option is impossible. But with the former alternative we get

$$T(\exists p)\,(Ms_1p \,\&\, {\sim}\,Tp)$$

i.e. $(\exists p)(Ms_1 p \,\&\, {\sim} Tp)$

so, say, $Ms_1 p_1 \,\&\, {\sim} Tp_1$

giving ${\sim} p_1$

but if $p_1 = (\exists p)(Ms_1 p \,\&\, {\sim} Tp)$

then ${\sim}(\exists p)(Ms_1 p \,\&\, {\sim} Tp)$

and so CONTRADICTION

hence $p_1 \neq (\exists p)(Ms_1 p \,\&\, {\sim} Tp)$.

This means s_1 is ambiguous.

So one wants to look at the following, which is a real 'liar', since it says not just that it is false, but also that it is not ambiguous:

This sentence expresses just one proposition and that is false.

In connection with this we can say

$$Ms_2 (\exists p)(Ms_2 p \,\&\, (q)(Ms_2 q \rightarrow p = q) \,\&\, {\sim} p)$$

and hypothesize, first

$$(\exists p)(Ms_2 p \,\&\, (q)(Ms_2 q \rightarrow p = q) \,\&\, {\sim} p)$$

so, say, $Ms_2 p_2 \,\&\, (q)(Ms_2 q \rightarrow p_2 = q) \,\&\, {\sim} p_2$

giving $p_2 = (\exists p)(Ms_2 p \,\&\, (q)(Ms_2 q \rightarrow p = q) \,\&\, {\sim} p)$

and hence ${\sim} (\exists p)(Ms_2 p \,\&\, (q)(Ms_2 q \rightarrow p = q) \,\&\, {\sim} p)$

so CONTRADICTION

This means that option is impossible. But hypothesizing, secondly

$${\sim} (\exists p)(Ms_2 p \,\&\, (q)(Ms_2 q \rightarrow p = q) \,\&\, {\sim} p)$$

this is $(p)((Ms_2 p \,\&\, {\sim} p) \rightarrow {\sim} (q)(Ms_2 q \rightarrow p = q))$

which yields, adopting the above 'p_2' as a convenient abbreviation,

$$(Ms_2 p_2 \,\&\, {\sim} p_2) \rightarrow (\exists q)(Ms_2 q \,\&\, p_2 \neq q)$$

giving $(\exists q)(Ms_2 q \,\&\, p_2 \neq q)$

This proves that s_2 equally is ambiguous.

We have, therefore, removed one major snag in Prior's system (or rather, in Prior's gloss on his system) which prevented a clear view of its application, and the prospects of a language containing its own semantics might already appear to be more bright. But a lot more work needs to be done, in a lot more areas, before that goal can be achieved, and so I shall finish by showing how Prior's approach may be of assistance on a far wider scale than even he envisaged. Prior's approach clearly, in the end, resolves The Liar and associated paradoxes like The Preface: less obviously, but more directly, it resolves the paradoxes of Satisfaction, Application, and Denotation.

Prior's basic insight is a Wittgensteinian one (see p. 35): the meaning of a sentence is not an object, it is something generated when that sentence is used rather than mentioned (this enables Prior to escape the Paradox of Analysis, see pp. 53–56); in a like manner the meaning of a predicate is not that predicate, or any other object, it is something formed when that predicate is used. Now if 'Py' is a one place predicate, an Extensional Logician would get into trouble, writing the Satisfaction Axiom (Schema)

$$x \text{ satisfies } `Py' \longleftrightarrow Px$$

for upon considering the one-place predicate

$$y \text{ does not satisfy } y$$

he would have to say

$$x \text{ satisfies } \textit{that} \longleftrightarrow x \text{ does not satisfy } x$$

so *that* satisfies *that* ⟷ *that* does not satisfy *that.*

But an Intensional Logician would get into as many difficulties if he re-wrote the Satisfaction Axiom

$$x \text{ satisfies the concept of being a } P \longleftrightarrow Px.$$

For 'the concept of being something which does not satisfy itself' would again be paradoxical, if concepts were reified and the question consequently arose whether this 'thing' satisfied itself. But if the concept is formed simply when a linguistic predicate is used rather than mentioned, there is no trouble, for then concepts cannot be referred to and so cannot take the place of 'x' in the above axiom. 'x satisfies the concept of being a P' is therefore best written 'It is true about x that it is P', making clear why one cannot, grammatically, form 'y does not satisfy y', i.e. 'It is not true about y that y'.

There is nothing grammatically wrong with saying that what applies to an object is 'P', or that what denotes an object is 'the P'; what is wrong is thinking that this is saying that what is true about the object is that it is P, or the P. So, in particular, '"heterological" applies to "heterological"' and '"the least number not denotable in English in less than 100 letters" denotes n' do not mean what they are often taken to mean. Prior allows (p. 15):

> ... it is ... true that, although *breaking one's leg* is quite obviously not a relation between a person and any form of words whatever, and to break my leg is certainly not to break the phrase 'my leg', nevertheless, in breaking my leg I do, *ipso facto*, stand in *a* relation to this phrase, since in breaking my leg I break something which would normally be described in English by this phrase. Breaking one's leg, all the same, does not *consist in* standing in this relation to this phrase, and 'I have broken my leg' does

> not *mean* 'I have broken what would normally be called in English "my leg" '.

So 'it is called "*P*" ' cannot do what it has often been called upon to do, namely express what is expressed by 'it is *P*'. Moreover, in certain cases there is not even a material equivalence, as with 'my leg'. Given that 'heterological' is defined so that $(x)(x \text{ is } H \longleftrightarrow x \text{ is not called } x)$ – the common attempt at a second-order definition '(P)('P' is $H \longleftrightarrow$ 'P' is not P)' is ruled out on formal grounds – then ' "*H*" is *H*' is materially equivalent not to ' "*H*" is called "*H*" ' but ' "*H*" is not called "*H*" '. This equivalence is not a contradiction, and it is only a paradox if we forget irony and think that the relation between a word and what it means is always direct. Consideration of 'the least number not denotable in English in less than 100 letters' also shows that there is not a general equivalence between '*n* is the *P*' and '*n* is called "the *P*" '. There is, of course, no such equivalence if 'the P' lacks a reference: in Strawson's terminology 'the *P*' then cannot be used, in Russell's *Principia* terminology 'ιxPx' then cannot be introduced – it is not, as Russell's 'On Denoting' article suggests, that 'ιxPx' is invariably defined and $Q(\iota xPx) \longleftrightarrow (\exists x)(y)((Py \longleftrightarrow y = x) \,\&\, Qx)$.

In general, it seems that it is forgetting literary devices like *double entendres*, irony, and the whole procedure of fiction which makes natural language seem paradoxical. But the denotation case may remind us of another thing it took Russell a long while to remember: it is the *demonstration* relation which is the only semantic relation which is strict (see Prior p. 156f). Prior points out (pp. 60–1) that quotation-marks are demonstratives (and hence make 'logically proper names'), so it is iff *w* is '*P*' that, necessarily, *w* is called ' "*P*" '.

The University of Western Australia
Nedlands, W. Australia 6009

©B. H. SLATER 1986

'IF A AND B THEN A'

By A. J. DALE

IN his paper 'If A and B then A' E. J. Lowe argues that we should not endorse the schema which serves as its title.[1] In this note I shall argue (in I) that Lowe's misgivings about the schema rest on a mistaken view about the truth-conditions of 'if ... then' sentences and (in II) that his proposed modification to the schema does little to remove the import of the Lewis proof that anything whatever follows from a contradiction.

---

[1] E. J. Lowe, 'If A and B then A', *Analysis*, 45.2, March 1985, pp. 93–8.

I

The schema $(A \& B) \to A$,[2] if taken as a theorem schema in the presence of the transitivity schema $A \to B, B \to C \; - A \to C$, leads to the schema $A \to C \vdash (A \& B) \to C$ which Lowe considers unacceptable because of the standard apparent counter-examples to it: from 'if Smith strikes this match then it will light' we should not infer 'if Smith strikes this match and holds it under water then it will light'. He is thus led to consider the validity of the transitivity schema itself and discerns a *fallacy of equivocation* committed in such examples. In the case of 'if Smith strikes this match and holds it under water then Smith strikes this match' an entailment is expressed whereas in the case of 'if Smith strikes this match then it will light' another kind of condition is expressed. It is because two different kinds of condition are expressed that we get the supposed failure of transitivity: whenever we get the same kind of condition expressed we have a valid instance of the schema. The first difficulty with this account is its reliance on 'kinds of condition' since no criterion has been provided by Lowe for establishing identity of kind. Consider an adaptation of an example I once gave to illustrate the possible invalidity of the transitivity schema: if Smith strikes the match it will light; if it lights oxygen is present; therefore, if Smith strikes the match oxygen is present.[3] It is clear that neither of the premises is an entailment, so if this is a counter-example to transitivity Lowe should be able to discern different kinds of condition expressed in these premises. If the grounds for the assertion of the premises are investigated we might, I suppose, say different kinds of ground are appealed to: for the first premise the observation that Smith, match and surrounding environment are normal and for the second the fact that oxygen is required for combustion. It is open to Lowe to discern different kinds of conditions by appeal to the different kinds of ground that he may detect here but at least he owes us some account of how fine these distinctions of kind have to be to ensure correct application of the transitivity schema.

The appeal to grounds leads me to my second and more fundamental argument. How true is it that an 'if ... then' sentence expresses an entailment or some other specified condition? We may certainly have different types of ground for asserting 'if ... then' sentences but we should not mistake a difference in the type of ground appealed to for a difference in the truth-conditions of the

---

[2] Throughout this paper '$\to$' represents 'if ... then', '$\Rightarrow$' represents 'entailment' and '$\Diamond$' represents 'it is logically possible that'.

[3] A. J. Dale, 'The Transitivity of 'if ... then" ', *Logique et Analyse* 1972, pp. 439–41.

'if ... then' sentence itself.[4] Consider the following analogy: 'John is related to Peter' may be true on the ground that John is the brother of Peter or on the ground that John is the father of Peter. But 'John is related to Peter' does not express the proposition that John is the brother of Peter nor the proposition that John is the father of Peter. Again the transitivity schema 'A is related to B; B is related to C, therefore, A is related to C' does not commit any fallacy of equivocation should the first premise be true because A is B's father and the second be true because B is C's brother.

Similarly an 'if ... then' sentence may be true because of an entailment linking antecedent to consequent or because of some scientific law which so links them; but this should be no more lead us to suppose that 'if ... then' sentences equivocate than that 'John is related to Peter' equivocates between the various possible family kinships.

II

Lowe believes that changing the theorem schema $\vdash (A \,\&\, B) \Rightarrow A$ to the inference schema $\Diamond (A \,\&\, B) \vdash (A \,\&\, B) \Rightarrow A$ will 'undercut the notorious "proof" that a contradiction entails anything whatever, because the "proof" assumes that "A & $\sim$A" entails "A"'. But the schema $(A \,\&\, B) \Rightarrow A$ is beside the point if we consider the same proof which leads to the conclusion that anything whatever follows from a pair of contradictory sentences A and $\sim$ A. Surely this conclusion is equally counter-intuitive, if not as notorious!

The University,
Hull

© A. J. DALE

---

[4] The same mistake seems to underline Strawson's assertion that 'if ... then' sentences have sometimes 'misguided second-order uses'; P. F. Strawson, *Introduction to Logical Theory*, 1952 p. 36.

REPLY TO DALE

By E. J. LOWE

IN 'If A and B, then A' (ANALYSIS 45.2, March 1985, pp. 93–8) I argued, amongst other things, for the validity of the transitivity schema $A \rightarrow B,\ B \rightarrow C \vdash A \rightarrow C$, provided that the conditional connective '$\rightarrow$' is interpreted univocally throughout *either* as being expressive of an entailment relation between antecedent and consequent *or* as being expressive of the sort of relation discernible in a conditional like 'If Smith strikes that match, it will light', in which the antecedent does not entail the consequent. A. J. Dale has now offered a putative counter-example which cannot apparently be turned aside as involving equivocation on this score (see his 'If A and B, then A', ANALYSIS 46.2, March 1986). Form 'If Smith strikes that match, it will light', and 'If it lights, oxygen is present', the schema would have us infer 'If Smith strikes that match, oxygen is present'. No doubt this would be an *odd* conclusion to infer from these premises, but it is very much open to question whether a *fallacy* is involved. If the inference is not valid, it should be consistent to affirm the premises and yet deny the conclusion. But – at least according to any system of conditional logic (such as Stalnaker's) which accepts the principle of conditional excluded middle – the negation of the conclusion is equivalent to 'If Smith strikes that match, oxygen is *not* present'; and it is far from evident that one should be prepared to affirm this if one has already affirmed 'If Smith strikes that match, it will light'! (Even if one rejects conditional excluded middle, one may still consider the negation of the conclusion equivalent to something slightly weaker like 'If Smith strikes that match, oxygen *may* not be present', which is scarcely less troublesome.) The oddity which attaches to Dale's example is not, I suggest, logical but pragmatic. The point is that, normally, someone would consider himself in a position to affirm 'If Smith strikes that match, it will light' only if he *already had* reason to suppose that oxygen was present *whether or not* Smith was going to strike the match, i.e. only if he considered himself in a position to affirm both 'If Smith *does* strike that match, oxygen is present' and 'If Smith *does not* strike that match, oxygen is present'. For familiar 'Gricean' reasons, it would therefore be quite inappropriate for such a person to *assert* the conditional 'If Smith strikes that match, oxygen is present' as a conclusion drawn from the given premises; but it doesn't follow that he would be committing any *logical* fallacy in so doing. (It is worth emphasising, I think, that even a standard possible-worlds account of conditionals like Stalnaker's, which rejects the transitivity schema, would not represent Dale's example as exhibiting a failure of transitivity; Dale has more than just me to contend with!)

A more fundamental objection raised by Dale questions my suggestion that there are indeed different *kinds* of conditional, some expressive of an entailment relation between antecedent and consequent and some expressive of some other sort(s) of relation. He suspects (though in point of fact he does nothing to *show*) that I mistake a difference between the *grounds for affirming* conditionals for a difference in their *truth-conditions*. It can hardly be said, however, that my proposed pluralism *as such* is particularly extravagant or implausible, in view of the fact that it is already very widely accepted that not *all* 'If . . . , then . . . ' sentences in natural language can be made to submit to a single semantic analysis. Moreover, anyone who hopes as I do (see my 'A simplification of the logic of conditionals', *Notre Dame Journal of Formal Logic* 24, 3, July 1983, pp. 357–66) to explicate conditionals in terms of modal concepts is almost bound to discern semantic distinctions between conditionals reflecting the distinctions between different kinds of modality (logical, causal, epistemic); and it would be by reference to these latter distinctions that I would, given more space, attempt to answer Dale's challenge to provide identity-criteria for the different kinds of conditional. Finally, it seems clear that, at least in the case of a good many conditionals in natural language, the affirmation of a conditional is intimately related to, if not actually expressive of, an endorsement of some type of argument from its antecedent as a premise (perhaps amongst others tacitly assumed) to its consequent as a conclusion. (We often actually explicitly *say* 'If such-and-such, then *it follows that* so-and-so'.) However, this being so we should naturally *expect* to find semantic distinctions amongst conditionals corresponding to the various different types of argument (demonstrative, causal, probabilistic) that we customarily engage in. (Nor is this observation at variance with my previous remark concerning the link between conditionals and modal concepts, since modal concepts and types of argument are themselves intimately related in obvious ways.)

In part II of his response Dale points out that while my central thesis (that a conditional of the form 'If A and B, then A' need not be regarded as true where A and B are mutually incompatible) may undercut the notorious 'proof' that a contradiction entails anything whatever, it does nothing to undermine a result which he finds equally counter-intuitive, namely, that 'anything whatever follows from a pair of contradictory sentences A and ~A'. However, I do not consider it to be an insignificant difference between the two 'proofs' that the first carries what may be felt to be an unpalatable implication about the nature of entailment which is not carried by the second: namely, the implication that some *single* sentence or proposition entails anything whatever. Thus the second proof, even if valid, licences no unpalatable affirmation of a conditional with an arbitrary consequent, for the simple reason that it provides no antecedent for any such conditional. That valid reasoning may lead one to any conclusion whatever if one starts from mutually incompatible premises strikes me as not especially surprising; but that it should constrain one to regard as a logical truth a conditional with a contradictory antecedent and an arbitrary consequent strikes me as very peculiar indeed.

University of Durham,
50 Old Elvet,
Durham DH1 3HN

© E. J. LOWE 1986

TRANSPARENTLY FALSE: REPLY TO HARDIN

By D. K. BUCKNER

C. L. Hardin ('A Transparent Case for Subjectivism', ANALYSIS 45.2, March 1985, pp. 117-19) argues, against McGinn, that the impossibility of a transparent white can be explained in terms of an 'objectivist' account, and has nothing essentially to do with the structure of perceptual experience. His case, as I understand it, is that all white objects are non-selective[1] *diffusing* reflectors, whereas all clear/transparent objects are *non-diffusing* reflectors/refractors. By definition, no object can be both diffusing and non-diffusing. Therefore – in virtue of the physical constitution of white objects – a transparent white is impossible.

Hardin is of course correct when he points out that mirrors are not white because (unlike ermine coats or snowballs) they reflect light coherently and evenly (i.e. non-diffusely), and it happens that we reserve the term 'white' for diffusing but non-selective reflectors, and 'clear' for non-diffusing, non-selective reflectors. But he misses or ignores the problem of why we have no ordinary language term to denote the property of being either white *or* clear. We *do* have a scientific (optical) term for this, namely, 'non-selective'. Both snowballs and mirrors are non-selective reflectors, i.e. they share the property of reflecting all parts of the incident spectrum. We also have ordinary-language terms to cover both diffusing and non-diffusing *selective* reflectors, e.g. we predicate 'green' of both wine-bottles and vine-leaves. Why should ordinary language lack a term that covers both the diffusing and non-diffusing *only* in the case of non-selective reflections?

Surely the only explanation is 'phenomenological'. Wine-bottles and vine-leaves have an optical property in common: they both absorb a certain part of the incident spectrum. But we call them both 'green' because they visibly share a certain *quale.* Now, optically speaking, wine-bottles are to vine-leaves as water is to milk. Wine-bottles and water are both non-diffusing refractors, vine-leaves and milk are both diffusing reflectors and the property shared by water and milk is that of refracting/reflecting all of the incident spectrum, i.e. non-selectivity. If this objective symmetry were to have a subjective correlate in the way that such substances visibly appear to us, we should expect there to be a certain visual quale common to both water and milk. But what characterizes water, window-panes and the air is the complete absence of any identifiable visual quale: they are 'colourless' in the sense that they have neither chromatic nor *achromatic* colour, and can thus hardly share colour with milk or any other visible substance. Hence there is no ordinary-language correlate for 'non-selective'; nor can this fact be explained in 'objectivist' terms. Achromatic transparency is, if you like, a visual *zero*: the absence of any visual quale whatsoever.[2] The problem for objectivism is not to find an objective correlate for achromatic transparency, since there clearly is one: non-selective non-diffusing refraction/reflection. The problem is to explain how the objective correlate is a specific physical property, that of reflecting or refracting light in a certain way, whereas its subjective counterpart is the lack of any property at all.

---

[1] To avoid confusion, I adhere to the following convention: 'non-selective' for the property of reflecting or refracting *all* of the incident spectrum (as a white, but not a grey object might do) – as opposed to 'uniformly selective' for the property of reflecting all parts of the incident spectrum in equal proportion (as both white and grey objects do).

So I think that McGinn's general case – that some facts demand a phenomenological explanation – is perfectly secure. The fact that achromatic transparency occupies the place of a visual zero in our perceptual framework, and that there is thus no visible similarity between milk and water, is not one that admits of a purely optical explanation. The reason must lie – somewhere – in an account of the 'phenomenological laws governing how things seem', the laws that 'arise from the very nature of perceptual existence'.[3]

University of Bristol,
9, Woodland Road, Bristol BS8 1TB

© D. K. BUCKNER 1986

---

[2] It may be objected that black holds this position. Is it soot, then, rather than water, that lacks visible qualities? One would have thought not. Besides, it must still be explained why water and milk share optical, but not visible qualities.

[3] I am grateful to an anonymous referee for *Analysis* for uncovering some dangerous ambiguities in the original draft of this paper.

AE or EA?

By H. W. NOONAN

IN his recent book *The Subjective View* (Clarendon Press, Oxford, 1983) Colin McGinn suggests that, in view of its quantificational complexity, two interpretations are possible for Frege's statement in 'The Thought' (in *Philosophical Logic*, ed. P. F. Strawson, Oxford University Press, 1967): '... everyone is presented to himself in a particular and primitive way, in which he is presented to no one else', namely:

(i) $\forall x \exists m$ (x is presented with x in m & $\sim\Diamond\exists y$ (y is presented with x in m & $y \neq x$))

and

(ii) $\exists m \forall x$ (x is presented with x in m & $\sim \Diamond (\exists y)$(y is presented with x in m & $y \neq x$)).

The difference is that the first reading allows that different people are presented to themselves in different ways, whereas the second claims that there is a single mode of presentation common to all people – they all think of themselves in the same way when they use 'I'. McGinn labels this second claim 'the EA thesis' and the opposing claim that different people are presented to themselves in different ways 'the AE thesis' (note that the AE thesis is stronger than (i), which is entailed by (ii), i.e. the EA thesis).

Frege is customarily interpreted as maintaining the AE thesis in 'The Thought', but McGinn argues that there is no good reason for this. True, the AE thesis is the only interpretation which is compatible with the classical Fregean thesis that sense (= mode of presentation) determines reference, but that thesis is not explicitly defended in 'The Thought' and, McGinn argues, the intrinsic implausibility of the AE thesis provides a reason for supposing that it was rather the EA thesis that Frege was maintaining there.

But it is not only the thesis that sense determines reference which belongs to the classical Fregean theory of sense and reference: other component theses are (a) that the sense of a sentence is a thought and (b) that the only consituents of the sense of a sentence are the senses of its parts.

In what follows I argue that when these components of the classical theory of sense and reference are kept in mind it becomes highly implausible to interpret Frege as maintaining the EA thesis in 'The Thought'.

The basic reason for this is very simple. Throughout 'The Thought' Frege consistently equates thought and sentential sense, that is, he maintains thesis (a). But he also insists that the thought I express with a sentence containing 'I' (e.g. 'I am ill') is distinct from the thought you express with that sentence. It is his view then that the *sense* I express with 'I am ill' is distinct from the *sense* you express with that sentence. Together with the EA thesis this entails (unless senses of subsentential expressions are individuated more finely than modes of presentation) that the sense of the sentence 'I am ill', as uttered by me, is not composed solely of the sense of 'I', as uttered by me, and the sense of 'am ill'. The EA thesis thus entails that this sense either contains some other component sense or contains a component which is not a sense at all. Neither of these possibilities is consistent with thesis (b) of the classical theory, and it surely lacks any plausibility to suggest that Frege would have maintained either position when he wrote 'The Thought'.

What the EA thesis naturally goes along with is an insistence that not only unambiguous *subsentential* expressions, but also entire sentences composed of such expressions, have the same sense, whatever the context in which they are uttered. This in turn entails a distinction, in the case of indexical sentences, between sentential sense and thought – where the latter is thought of as possessing a context-independent truth-value. The difficulty for McGinn is just that Frege makes this distinction nowhere in 'The Thought' and writes throughout in a way that is inconsistent with it.

Of course, it is open to McGinn to argue that Frege *ought* to have espoused the EA, rather than the AE, thesis in 'The Thought', and that he *ought* to have distinguished, in the case of indexical sentences, between thoughts *qua* bearers of absolute truth-values and sentential senses. But the fact is that the text of 'The Thought' makes it absolutely clear that Frege did not do the latter of these things, and thereby makes it exceedingly plausible that he did not do the former either.

University of Birmingham,
P.O. Box 363,
Birmingham B15 2TT

© H. W. NOONAN 1986

HUME'S ARGUMENT FOR THE TEMPORAL PRIORITY OF CAUSE OVER EFFECT

By MICHAEL J. COSTA

THE following passage in Hume's *Treatise* [1] has proved to be puzzling for a number of commentators. In it Hume presents an argument for the temporal priority of cause over effect.

> The second relation I shall observe as essential to causes and effects, is not so universally acknowledg'd, but is liable to some controversy. 'Tis that of priority of time in the cause before the effect. Some pretend that 'tis not absolutely necessary a cause shou'd precede its effect; but that any object or action, in the very first moment of its existence, may exert its productive quality, and give rise to another object or action, perfectly co-temporary with itself. But beside that experience in most instances seems to contradict this opinion, we may establish the relation of priority by a kind of inference or reasoning. 'Tis an establish'd maxim both in natural and moral philosophy, that an object, which exists for any time in its full perfection without producing another, is not its sole cause; but is assisted by some other principle, which pushes it from its state of inactivity, and makes it exert that energy, of which it was secretly possest. Now if any cause may be perfectly co-temporary with its effect, 'tis certain, according to this maxim, that they must all of them be so; since any one of them, which retards its operation for a single moment,

> exerts not itself at that very individual time, in which it might have operated and therefore is no proper cause. The consequence of this wou'd be no less than the destruction of that succession of causes, which we observe in the world; and indeed, the utter annihilation of time. For if one cause were co-temporary with its effect, and this effect with *its* effect, and so on, 'tis plain there wou'd be no such thing as succession, and all objects must be co-existent. (T. 75–6)

Barry Stroud finds the argument 'extremely puzzling' because its conclusion seems to contradict one of its premises ([2], pp. 253–4). Beauchamp and Rosenberg say that the argument is 'atypical of Hume, for it is both obscure and ill-arranged' ([3], p. 192). One must admit that the argument is a bit obscure, but I shall try to show that, properly understood, it is not ill-arranged and does not commit any obvious logical blunders.

It is rather clear what Hume takes the argument to establish. He intends to answer those who 'pretend that 'tis not absolutely necessary a cause shou'd precede its effect,' so the conclusion of the argument must be that it *is* absolutely necessary that a cause should precede its effect. Let us call this conclusion the 'Causal Temporal Priority Thesis' or 'CTP' for short. The argument appears to be of the form of a *reductio.* The hypothesis that is to be reduced to absurdity is that 'any cause may be perfectly co-temporary with its effect'. The other explicit premise is the "maxim" that 'an object which exists for any time in its full perfection without producing another, is not its sole cause'. It is said to follow from these premises that all causes must be perfectly contemporary with their effects. And it is said to follow from this that there is no succession of causes or time in the world. Hume says that we can observe that this consequence is false. Therefore, we may conclude by *reductio* that CTP is true.

The most obvious way to read the *reductio* hypothesis is as the existential statement 'Some cause is perfectly contemporary with its effect'. The denial of this statement (together with the implicit supposition that causes cannot follow effects) does entail that all causes precede their effects. Reading the hypothesis in this way, we get the following explicit representation of Hume's argument.

RH: Some (proper) cause is perfectly contemporary with its effect.

MAXIM: Any cause which retards its operation for any time is not a sole (or proper) cause.

Therefore,

IC: All (proper) causes are perfectly contemporary with their effects.

CP: If all (proper) causes are perfectly contemporary with their effects, there is no causal succession or time.

Therefore,

RC: There is no causal succession or time.

But, we observe that

RP: There is causal succession and time.

Therefore (by *reductio*),

CTP: A (proper) cause must temporally precede its effect.

No doubt there are interesting philosophical issues involved in determining whether MAXIM and CP are true, but even if we accept MAXIM and CP problems remain. First, we can see the basis for Stroud's complaint that the conclusion of the argument contradicts one of its premises. MAXIM seems to preclude the possibility that a cause precedes its effect. Second, it is not at all obvious what work RH does in producing the intermediate conclusion IC. It seems that IC follows from the second premise, MAXIM, alone. If this is the case, then the conclusion, by *reductio*, should be the denial of MAXIM. (CTP, indeed, would follow validly in the argument indicated; but only because anything can be deduced from a contradiction.) Moreover, Hume characterizes MAXIM as 'an establish'd maxim;' and he gives no indication that he takes the argument to shed any doubt whatsoever on it.

Beauchamp and Rosenberg reconstruct the argument in the form of a dilemma inside of a *reductio*, which has the hypothesis 'One or more causes are contemporaneous' ([3], pp. 193–4). So far as I can see their reconstruction has all of the problems identified above in addition to the drawback of maintaining that Hume seriously misrepresented the structure of his argument.

I think I have found a way to avoid these problems. The key is to suggest a new way to read the *reductio* hypothesis of the argument. Hume's own way of expressing this hypothesis is 'any cause may be perfectly co-temporary with its effect'. I propose that we take the 'may' in Hume's wording seriously. In lieu of RH, let us read the first premise as PRH: It is possible that causes are perfectly contemporary with their effects. If PRH, then MAXIM requires in fact that all causes be perfectly contemporary with their effects; but this produces RC and, by *reductio*, the denial of PRH, viz., it is not possible that causes are perfectly contemporary with their effects. The denial of PRH (together with the implicit supposition that causes cannot follow effects) entails CTP. We get the desired conclusion through an argument that matches point by point the form suggested by Hume's own words. All that we need to do is to replace RH with PRH.

Moreover, we do not need to reject MAXIM! If it is impossible that a cause be perfectly contemporary with its effect, then MAXIM cannot be read as demanding this. But a careful reading of MAXIM shows that it requires only that there be no interval between the proper cause and its effect. Suppose you held the view that time consisted in discrete, durationless instants. And suppose you believed that events could occupy such instants. Then MAXIM could be true without (per impossible, per Hume) causes being perfectly contemporary with their effects. Each effect has its proper cause occurring in the immediately preceding instant. That cause does not exist for any *time* without producing its effect (there is no time or duration until you have a succession of events). This is not an idle supposition, of course, for Hume explicitly endorses and argues for just such an account of time in Book I, Part II of the *Treatise*.

Does it make sense to think that an event might exist in a durationless instant? I think that for Hume, at least, it does. Hume's argument for the nature of time is patterned after a corresponding argument for the nature of space as consisting of extensionless, indivisible points. Hume saw no absurdity in holding that an extensionless point is coloured, so there is no reason to think that he would see anything absurd in holding that a durationless instant might contain an event. Indeed, Hume seems to think that instants *must* contain events. He says, '... the indivisible moments of time must be filled with some real object or existence, whose succession forms the duration, and makes it be conceivable by the mind' (T.39).

None of this commits Hume to the view that the events we commonly talk about are durationless. An event having duration may be composed of a sequence of instantaneous, durationless events (just as a surface having spatial dimensions would, for Hume, consist of an array of extensionless points). All that is required is that *proper* causes be instantaneous events. When we loosely, in day to day affairs, identify an event as a cause of another, we typically identify as cause an event that merely contains the proper cause of an event (the proper effect?) that is a part of the event comprising the effect.

Such a view of time and events may seem hopelessly outmoded (although it is not clear to me that contemporary accounts of time and events are devoid of nonsense), but there is no reason to believe that Hume would have seen it as absurd. It is perfectly reasonable to think that Hume had his view of time and events in mind as he formulated his argument for CTP.

University of South Carolina,
Columbia, S.C. 29208, U.S.A.

© Michael J. Costa 1986

References

[1] David Hume, *A Treatise of Human Nature*, ed. L. A. Selby Bigge, Oxford 1888.
[2] Barry Stroud, *Hume*, Routledge & Kegan Paul, 1977.
[3] Tom Beauchamp and Alexander Rosenberg, *Hume and the Problem of Causation*, Oxford, 1981.

THE IMMATERIALITY OF 'ABSTRACT OBJECTS' AND THE MENTAL

By IRVING THALBERG

SUPPOSE each one of us somehow 'has' or 'is' in a mind, which shelters various entities or events: pains or bouts of discomfort; thoughts or episodes of pondering; desires or surges of longing. Mind-body dualists consider these all to be non-physical. Materialists demur. Yet few debaters on either side have asked what these mental items or events would be like if they *were* incorporeal. The scant information we've culled sounds irreducibly negative. Our mind and its denizens are said to have *no* location or other spatial attributes, to be *in*divisible, *not* accessible to or ever shared with another person, *not* perceptible by means of our sense organs, but known *in*fallibly, *non*-inferentially, *without* evidence, and to *no one* but ourselves.

Two October 1984 journal articles address this neglected topic. Writing in *American Philosophical Quarterly* (xxi, pp. 367-70) G. J. Reid complains:

> The ... materialist attack upon the category of the immaterial ... has taken the form of claiming that we have no need of [it] ... whatever it might amount to in detail ...
>
> The materialists ... have not been much concerned to articulate [it before excluding it].

More significantly, Reid asserts that 'the immaterial' is not doomed to be 'negatively defined (as the non-extended, for example)'. In fact Reid has 'three positive accounts' of it.

Before we plunge into that, we should take note of the other recent discussion, in *Mind* (xciii, pp. 556-9). Its author, Peter Glassen, seems happy with a negative account. He says

> we ... mean by 'non-physical' ... 'lacking one or more of the defining characteristics of the physical' ... [F]or our ... purposes ... 'physical' objects and events are [defined as] those that *occupy space*, that are *publicly observable*, and that *may enter into causal relations* ...

This proposal generates fascinating side-issues. First, *could* any object or event lack just 'one ... of the defining characteristics of the physical'? For instance, might it lack a habitat, yet be observable – and also figure in causal transactions? Alternatively, could something be unaffected by its surroundings, and causally powerless itself – hence incapable of stimulating our senses – but publicly perceptible? A second puzzle is that even if such conglomerations of present and absent 'defining' traits of 'the physical' occur, will they always rank as non-physical?

Glassen would avoid both difficulties if he required that non-physical things lack *all* the characteristics he mentions. But that will create a third problem: won't interactionist dualists flatly *deny* that mental – and by hypothesis non-physical – events *lack* the characteristic of being able to affect and be affected by physical happenings? Some parallelists and epiphenomenalists will object too, that at least *mental* occurrences interact.

There is a similar hitch about lacking the characteristic of extension: how should we classify the visual images, pictorial 'mental representations' and sense data which many dualists say we inwardly contemplate? Allegedly these items are located *within* our mental sanctuary, and have shape, colour and relative size – all of which involves space. How then can they, and the mind housing them, qualify as immaterial? Glassen displays some cognizance of this fourth puzzle. He admits it is a mystery 'what could be going on in something... that does not occupy space'. But he reassures us that it is *meaningful* to say 'what goes on [there is] imaging and perceiving'. And generally a 'dualist's claim that the mental is non-physical in that it lacks ... [the characteristic] of occupying space, is perfectly meaningful'.

Now forget my misgivings towards Glassen's analysis of what 'we ... mean by "non-physical"'. Assume that we *can* meaningfully speak of extensionless minds and of the imaging that takes place within them. How has Glassen enlightened us about the immateriality of mental objects and events? We had found it uninformative to be told what immaterial things are *not*. Are we better off hearing which attributes they *lack*? Someone should reveal what attributes are *left*. I doubt that the central question is 'What do terms like "immaterial" and "extensionless" *mean*?', or 'Are they meaningful?' Instead we should ask: What are *any* features – 'defining', normal, accidental, even idiosyncratic – that a bodiless thing might exhibit? Naturally if a dualist answers by ascribing to a non-physical item features which define or typify 'the mental', that will make her or his claim, 'the mental is non-physical', an empty tautology.

If we were inquiring what everyday negatively described things are like, would our task be so arduous? Glassen thinks it would, and offers three down-to-earth examples. He says that apparently

> [a] 'non-prescription' remedy can be *defined* in positive terms as one that one is legally permitted to buy 'over the counter' ... [T]his seems to be a *definition* in positive terms; but is it really? Doesn't it *just mean* 'without a doctor's written authorization ... '? If it *means* anything other than this, I don't know what it is . [Similarly, an] 'immobile' object is one that 'stands still' ... But what does 'stands still' *mean*? The *only sense* I can make of [it] is: 'does not change its position ... ' [Again,] I list the ingredients of a number of refreshments. What do I *mean* when I say ... these refreshments are non-alcoholic? Why surely that not one ingredient is an alcohol (some italics added).

Once more I doubt that the issue is whether 'non-physical' – or humbler negative expressions – are meaningful. Even *what they* *mean*, strictly speaking, and how we should narrowly *define* them, may be irrelevant. In this regard, 'non-prescription' and its workaday kin are on a par with 'non-physical' and its noble cognates. Maybe we should only define, or articulate the exact *sense* of, both species of negative expression in further negative phraseology. Yet a vital difference remains. We *can* portray, in lavish detail, various non-prescription medicines, immobile house-hold items such as bathtubs, and familiar non-spirituous beverages like V8 juice. By contrast, when Glassen and others tell us that dualists claim 'the mental lacks defining characteristics of the physical' – above all the 'characteristic of occupying space' – our progress is minimal. What else is there for the mental to do?

For his part, Glassen thinks we are not really in the dark about what it is for 'the mind and its events' to be bodiless and 'not occupy or occur in space'. He argues that *numbers* are clearly immaterial entities, and that we 'understand . . . what it would mean to say that a number had no spatial position'; we should moreover agree 'on the necessity of [numbers] not occupying space'. In response I'm tempted to invoke Wittgenstein's comparison of two negative statements, 'A goose has no teeth' and 'A rose has no teeth' (PI, pp. 221f.). Isn't the second statement necessarily or 'obviously true', even 'surer' than the first? Wittgenstein decides this is 'not so clear. For where should a rose's teeth have been? The goose has none in its jaw. And neither . . . has it any in its wings; but no one means [or asserts] that when he says it has no teeth'. By analogy, I am dumbfounded as to how things would be if some number – or all numbers – had spatial attributes. Consequently I'm altogether uncertain what we are declaring impossible when we say they have 'no spatial position', and call this a 'necessity'.

Suppose I'm deluded, and relatively rational inquirers like myself can understand what it is for *bona fide* non-physical entities of one sort (numbers) to lack extension. Will this serve as a model for clarifying the putative non-spatial and non-physical character of 'the mind and its events'?

Reid's 'three positive accounts of the immaterial' should come in handy at this juncture. They are, nearly *verbatim*:

(i) 'Whatever can be shared without diminishment is immaterial'. In other words: 'whatever can be multiplied without loss of identity' or singleness will be non-physical. Reid says 'abstract objects', including numbers and 'universals' like 'a shade of red', all pass this test; 'trees and stones' do not. Reid further claims that 'knowledge is immaterial', perhaps failing to discriminate between (A) *what* several researchers at different times know to be true – which might be a single 'abstract' proposition – and (B) the myriad of events, none of them 'abstract', which constitute each person's acquiring knowledge of it.

(ii) 'Whatever is intentional is immaterial'. Reid explicates 'intentionality' by distinguishing now 'between content and occurrence'.

The content of an occurrence would be what the occurrence of 'of' or 'about'. My example would be the proposition which is what a researcher's knowledge is knowledge of. Reid's illustrations are '[a]n impression of red' and 'statues of something'. Reid asserts that 'statues [of something] are intentional', but wisely neglects to combine this with the major premise above, 'Whatever is intentional is immaterial'. Of course the premise should have been: whatever is intentional has an abstract *content* which is not a material object or event. Then neither a stone sculpture nor an 'impression of red' must be incorporeal.

(iii) 'Immaterial items, unlike material items, have clean cut identity features'. A single 'daffodil bulb' can split 'into two ... each of which has an equally good claim to be [identical with] the previous [*sic*] bulb'. Yet 'whatever empirical multiplicity we ... encounter cannot be explained as two different ... number 9s'. I suspect all this was implied by account (i) of 'whatever can be multiplied without loss of identity'.

At least Reid's (i)-(iii) furnish positive information about one species of 'the immaterial'. All the non-bodily items Reid surveys are 'abstract objects': numbers, universals, contents of sensory 'impressions' and other mental states, along with contents of representational artefacts. But Reid has not shown or illustrated how psychological events – much less our 'statutes of something' – might be incorporeal too. Thus Reid's account (i) misclassified the mental state of knowing. The mixup becomes glaringly evident when Reid announces that 'items ... as common ... as ... books ... are ... immaterial'. Isn't a book's content a likelier candidate? I react similarly to Reid's insinuation that 'consciousness' is 'not material' by (i)-(iii). In fact (i)-(iii) only provide that contents of awareness – what we are conscious of – are abstract, reproducible yet one, and so non-physical.

(i)-(iii) make it obvious that the characteristics which endow abstract objects with their non-physical standing cannot belong to minds and their states. According to almost any conception of the mental – dualist, materialist, whatever – there is no analogue to the sharing and multiplication 'without diminishment' that typifies abstract objects. By the same token abstract items seem not to be the right sort of things to have, *or* to *lack*, features which distinguish the mental. Your mind, and your state of expecting a rainstorm, must be uniquely yours; they begin and cease to exist in time, and undergo alterations; many theorists say they are introspectible, perhaps unerringly. How could a number be owned or unowned, have or not have a beginning and ending, change or resist change? A number seems unsuited for introspection, accurate or inaccurate, and also for public observability or invisibility.

But isn't the mental, as Glassen said, like numbers and other abstract entities in being deprived of extension? I think the analogy is dubious. Perhaps numbers are as ill suited to be *lacking* an abode in space as to be a tenant. They may be of the wrong category to have extension *and* to be extensionless.

This analysis seems not to carry over to minds and psychological happenings. Materialists from antiquity to the present have sometimes identified our minds with a bodily organ. Once the heart was a favourite, but for centuries now the brain has taken over. Most often, materialistically inclined theorists have equated mental events – such as episodes of musing, processes of deliberation, and outbursts of hostility – with a person's relevant behaviour and whatever neural events are needed. No doubt publicly recognized practices and rules must also be in the background. However, this crude sketch is enough to indicate that minds and psychological states, unlike numbers, are at least *eligible* to have or to lack extension.

Of course dualists and their interpreters, including Reid and Glassen, have not told us how things would be if the mental were extensionless and therefore non-physical. Yet that is secondary. What strikes me is this. We have moderately positive accounts of abstract entities and the characteristics which differentiate them from every material particular. But because they and the mental are so heterogenous, our study of their immateriality has failed to advance our understanding of a dualist's war cry, 'the mental is non-physical'.

University of Illinois at Chicago,
Box 4348, Chicago, Illinois 60680, U.S.A.

©IRVING THALBERG 1986

CHURCHLAND ON DIRECT INTROSPECTION OF BRAIN STATES

By NATIKA NEWTON

PAUL Churchland has recently argued that the phenomenal features of our experiences are not a barrier to psycho-physical reduction (P. Churchland, 'Reduction, qualia and the direct introspection of brain states', *Journal of Philosophy*, February 1985, pp. 2-28). His position is that introspection of qualia is introspection of brain states; with the conceptual framework provided by a mature neuroscience we can learn to achieve 'direct, self-conscious introspection' of such properties as spiking frequencies of neural pathways.

I am in sympathy with Churchland's general goal of showing that phenomenal qualities of experiences can be accounted for in a physicalist philosophy, and I am also in sympathy with his claims that sophistication in neuro-science may well allow us to make finer, more accurate, and more useful discriminations in our sensations. But I disagree with his claim about the direct introspection of brain states. Churchland appears to think of introspection of our sensations and their properties by 'inner sense' as analogous to the perception of external objects and their properties by 'outer sense'. This is the approach that Sydney Shoemaker has called the 'perceptual model' and criticized extensively (S. Shoemaker, *Self-knowledge and self-identity*, Cornell 1963; and 'Self-reference and self-awareness', *Journal of Philosophy*, October 1968, pp. 555-67). I want to show that there is a crucial disanalogy between perception of objects and introspection of sensations, such that *even if* a perceptual model were justified, Churchland's conclusions in this case do not follow. Specifically, even if we do introspect our sensations analogously to the way we perceive objects, it may be that we necessarily introspect them *as sensations*, rather than as brain states, even if that is what they are. As a matter of fact, I believe that there are difficulties with the notion of introspecting sensations *as sensations*, but I shall not discuss that in this paper.

Let us assume that it is correct to say that there is introspective awareness of qualia of sensations. What does this mean? Sensations are internal states associated with the appearance of objects to us. Properties of the sensation, qualia, are features of the appearance of the object. In perceiving an object we can pay attention to it, or to the way it appears to us — the sensation and its properties. If we pay attention to the sensation instead of the object, then certain properties, such as colour or taste, remain "in view", while others, such as inferred relations between this and other absent objects, are ignored. Paying attention in this way to the sensation is having introspective awareness of the qualia of the sensation.

Churchland holds that sensations are reducible to and identical with brain states, that the properties of sensations are identical with and reducible to properties of brain states and that introspecting, say, the searing quality of my pain is introspecting the spiking frequency of my C-fibres. Our current lack of neuro-scientific sophistication results in introspective awareness being 'opaque': I can think about the searing quality without having the least idea of the spiking frequency, and I might even deny that they are or could be identical. With enough knowledge of neuroscience, however, I can learn to experience qualia *as* brain properties, not just as phenomenal properties of my sensations.

Churchland's claim seems to be based on an assumption (with which I am not quarrelling) about perception: that if you perceive x, and $x = y$, then you are perceiving y (perceiving is extensional); furthermore, if you perceive x as having the property F, and you learn that $F = G$, then you can perceive x as having the property G. In effect, you can now reduce the property F to the property G.

My discussion centres on two questions:

(a) what does it mean to say that 'one perceives x as *G*?'
(b) Can the possibility be realized in the case of sensations and brain events?

With regard to the first, let us consider an example offered by Churchland of the sort of perceptual change he has in mind. He describes a child who first hears the sound of a symphony orchestra (*x*) as a seamless voice (*F*) and later learns that the orchestra consists of individual instruments (*y*) playing distinguishable and identifiable chords supporting a melody line (*G*). His claim is that neuroscientific sophistication will do for introspection of sensations what musicology will do for hearing music: we will learn to introspect our sensations *as* brain states (e.g. dopamine levels) in the same way that the sophisticated musicologist has learned to hear the sound of the orchestra as Gm7 chords and Adim chords.

It is not completely clear what Churchland has in mind when he talks about the change that neuroscientific sophistication might bring. There are two possibilities, and each is problematic. On the one hand, he might mean *just* that we can learn to use neuroscientific language in our reports instead of phenomenal language:

> Suppose we trained our native mechanisms to make a new and more detailed set of discrimination, a set that corresponded not to the primitive psychological taxonomy of ordinary language, but to some more penetrating taxonomy of states drawn from a complete neuroscience. And suppose we trained ourselves to respond to that reconfigured discriminative activity with judgements that were framed, as a matter of course, in the appropriate concepts from neuro-science. (Churchland p. 16).

In this case he is not using a perceptual model. We might get so that we always automatically describe our sensations in neuroscientific terms. If we did this accurately and with finer discriminations, the result could be a big improvement in self-knowledge. But it doesn't seem right to say that this would be a change in the *way we introspect* our sensations, such that we then introspect them *as* brain states. It seems as though someone could learn to talk about spiking frequencies when he feels pain, while still feeling the pain *as pain*, with all of its phenomenal properties. In any case, Churchland clearly implies that we can do more than this. He says

> Consider now the possibility of learning to describe, conceive, *and introspectively apprehend* (my italics) the teeming intricacies of our inner lives within the conceptual framework of a matured neuroscience ... the enhancement of our introspective vision could approximate a revelation. Dopamine levels in the limbic system ... could be moved into the objective focus of our introspective discrimination ... [this change will constitute] not a gloom in which our inner life is suppressed or eclipsed, but rather a dawning, in which its marvelous intricacies are finally *revealed* – most

> notably, if we apply ourselves, in direct self-conscious introspection. (Churchland, pp. 16, 28).

Churchland seems to have in mind not just a change in reporting ability but an experiential change. I do not only learn that the property *F* of the *x* I have been 'apprehending' is identical with the property *G*, but I learn to apprehend *x* *as* *G* instead of or in addition to *F*.

If Churchland means that we might come to introspect our sensations as brain states in the sense that we could learn to have a new experience, then he is using a perceptual model and there is a different objection.

There is a crucial disanalogy between learning to perceive the sound of the orchestra as a collection of identifiable chords, and learning to introspect sensations as brain states. The latter involves an aspect shift that the former does not. Hearing the chords in the orchestra just is making finer discriminations and having some new language to describe them in, tasting the chemicals in the wine is learning to give chemical names to the qualities one has learned to pick out in the sour taste one was already familar with. In both cases one learns to discriminate more detail in the original object of perception. But in the case of introspection, there would be an abrupt change. Not only does pain normally not appear as properties of neurons, but it appears as something quite other: in particular, it usually does not seem to be in one's head. I can imagine learning to make detailed discriminations in that throbbing sensation in my foot which would match in precision the sorts that Churchland has in mind, and *also* learn to use neurological language to refer to these, while still experiencing the pain *as* a throbbing sensation in my foot. But Churchland seems to want something else. We are to stop experiencing pain as something in our foot, and instead experience it as something in our head.

The two types of case might be analogous to two changes that might occur in the way one views Wittgenstein's duck-rabbit drawing illustrating aspect shift. The orchestra example corresponds to a shift between seeing the drawing naively just as a duck, and sophisticatedly as, say, a Peking duck in an aggressive stance. But the sensation case corresponds to a shift between naively seeing the drawing as a duck and sophisticatedly seeing it as a rabbit. But if Churchland does have this sort of shift in mind, he begs a question: how do we know that sensations are capable of appearing both phenomenally *and* as brain properties, in the way that the drawing is capable of appearing as a duck or as a rabbit? The drawing was constructed expressly to make this perceptual shift possible, but we have been given no reason to think that sensation qualia are such as to be introspectable *as* brain properties. So neither the orchestra example nor the wine-tasting example support the claim that this sort of shift is possible.

One can object to my interpretation of the examples of external perception. It might be thought that aspect shifts *are* involved in the types of cases illustrated by the orchestra example, and hence that they are not so different from the introspection case as I am claiming. Suppose someone learns that the slimy green surface on the pond water he has been seeing consists of innumerable algae, massed together. Isn't it an aspect shift to think of many algae instead of solid green slime (I am indebted to Christopher Hill for this objection)? The answer, I think, depends on the way one 'thinks of' the many algae. If one thinks of them as massed together and seen by the naked human eye, then one is just thinking of them in terms of their appearance *as* green slime, and there is reduction and no aspect shift. On the other hand, one could think of the algae by visualizing individuals or small groups of them. In this case there would be an aspect shift; but here the example is no longer relevant for Churchland's purposes. In order to visualize individual algae I must stop paying attention to the appearance of the green slime. *That* phenomenon is not part of my thought about the algae; I can think of them just as well (or better) without it. Distinctions I see among parts of the slime are not reducible to distinctions I visualize among individual algae. So on one reading we see green slime as algae just by describing it that way; that means we can reduce it to algae but only by a change in reporting ability, which as I have argued does not correspond to learning to introspect sensations *as* brain states. On the other reading we learn to substitute a *visualization* of algae for the visual appearance of slime, or else we superimpose the one on the other. This latter reading is indeed analogous to what would probably happen with the neuroscientific sophistication Churchland envisages. We would learn to think about brain properties while experiencing the phenomenal properties of sensations. But this would not be introspecting sensations *as* brain properties.

But it must be noted that, while this objection does not work, some examples Churchland gives of external perceptual changes do involve aspect shifts of the sort he seems to have in mind for introspection. He speaks of the astronomer 'for whom the speckled black dome of her youth has become visible abyss' (Churchland p. 15); other examples he has given include coming to see the earth revolving around the sun, or coming to see witches as harmless old women (P. Churchland, *Matter and consciousness*, MIT Press 1979, p. 44). But these are cases of *eliminativism* rather than reductionism; here a sudden aspect shift makes sense. We don't see a black dome as an abyss while continuing to see it as a dome; the two are mutually exclusive, and a new theory has replaced an old one. These examples seem much closer to what Churchland seems in some places to be predicting for introspection if the shift he describes were to happen; on the other hand, his arguments that it could happen depend on the *reductionist* examples of the symphonygoer and the wine-taster. So he has not established that we could learn to introspect sensations *as* brain properties.[1]

C.W. Post Center,
Long Island University,
Greenvale, New York 11548, U.S.A.

©NATIKA NEWTON 1986

---

[1] Much of the work for this paper was done during a National Endowment for the Humanities Summer Seminar at Cornell University in 1985. The author is grateful to its director, Sydney Shoemaker, and to the other seminar participants, for helpful advice and criticism.

NOZICK, JUSTICE, AND THE SORITES

By ROY A. SORENSEN

CENTRAL to Robert Nozick's entitlement theory of distributive justice (*Anarchy, State and Utopia*, Blackwell, 1974) is the following principle:

> N 'Whatever arises from a just situation [distribution] by just steps is itself just' (p. 151).

Edward Quest in a note titled after N (ANALYSIS 37.4, June 1977) has objected to the principle by pointing out its asssociaton with the sorites paradox. David Lederkramer has defended Nozick's appeal to N by arguing that distributive justice is not inexact under the entitlement theory and by raising other objections to Quest's application of the sorites problem ('Quest on the Entitlement Theory', ANALYSIS 39.4, October 1979). Although no defence will be provided of Quest's version of the sorites objection, I shall present another version which preserves Quest's insight that the sorites poses a serious problem for Nozick's appeals to N.

Since nearly all of our vocabulary is vague, Lederkramer's claim that 'justice' is not vague is *prima vacie* implausible. Of course this surface implausibility would diminish if we learned that Nozick employed a precisified definition of 'justice'. However, Nozick prefaces his book with the following disclaimer:

> This book does not present a precise theory of the moral basis of individual rights; it does not contain a precise statement and justification of a theory of retributive punishment; or a precise statement of the principles of the tripartite theory of distributive justice it presents. (p. xiv)

Rather than appealing to an antecedent precisification of 'justice', Lederkramer bases his claim on a criterion of vagueness. According to Lederkramer, a predicate is vague just in case its threshold for application cannot be nonarbitrarily specified. For example, 'heavy' is vague because any statement such as 'x is heavy if x weighs over 50 pounds' would display some arbitrariness. Although the criterion may be ill suited to accommodate conflict vagueness, it is fairly typical and unobjectionable. Lederkramer continues:

> We have, then, a criterion for distinguishing inexact predicates from others; inexact predicates are those which are ascribed according to the previously noted rule. The predicate 'is entitled to' is not, therefore, inexact. Thus, any theory of distributive justice which stipulates that 'a distribution is just if an only if everyone is entitled to his holdings' is not an inexact theory; that is, under such a theory the predicate 'is just distribution is not inexact. (p. 221).

Lederkramer's reasoning appears to be: 'just distribution' is vague under the entitlement theory only if 'entitled' is vague; but 'entitled' is not vague because it is free of arbitrary thresholds; so Nozick's theory of distributive of justice is free of sorites difficulties.

Although there are no obviously arbitrary thresholds for 'entitled', as there are for 'heavy', 'poor', and 'heap', the fact that sorites have been constructed for 'stone', 'table', 'person', and even 'vague' encourages the search for subtler thresholds. A promising area of investigation is Nozick's discussion of the Lockean proviso. Nozick interprets the proviso as a kind of catastrophe clause which prevents individuals from cornering the market on goods essential to survival.

> If the proviso excludes someone's appropriating all the drinkable water in the world, it also excludes his purchasing it all. (More weakly, and messily, it may exclude his charging certain prices for some of his supply.) ... His later acquisition of it all does not show that original appropriation violated the proviso ... Rather, it is the combination of the original appropriation *plus* all the later transfers and actions that violates the Lockean proviso. (pp. 179–180).

As Nozick puts it, 'the historical shadow of the Lockean proviso on appropriation' is included in every owner's title to his holdings. Whether I corner the market deliberately, through a clever chain of transactions, or inadvertently, through a natural disaster which destroys the supplies to my competitors, the Lockean proviso is activated and severely restricts what I can do with my 'property'. However, there can be unclarity as to whether the Lockean proviso applies. Although the proviso is activated if I buy up all the water, it is also activated if I buy all the water with the exception of one drop. Likewise it is activated if I leave only two or three drops for others. What is the minimum number of drops I can leave others to ensure that the Lockean proviso does not apply to me? Since we cannot answer this question nonarbitrarily, it follows that 'falls under the Lockean proviso' is a vague predicate. The vagueness of this predicate infects Nozick's 'entitled' predicate because what I am entitled to depends on whether the Lockean proviso applies. The vagueness of 'entitled' in turn infects 'just distribution' since 'a distribution is just if and only if everyone is entitled to his holdings'. And it is the vagueness of 'just distribution' which is the seed of the sorites difficulties.

For example, suppose there are one hundred individuals and one billion litres of water. Originally everyone appropriates equal shares of the water supply. This is a just distribution. (Those disagreeing can substitute another distribution and make adjustments accordingly.) Justice is preserved if I purchase no water from my neighbours because it leaves the original distribution intact. Justice is also preserved if I purchase 1 litre or 2 litres. Any one litre purchase is insignificant. We can now formulate a sorites argument to show that it is just of me to corner the market.

(1) Base step: The water distribution is just if I purchase 0 litres of water.
(2) Induction step: If the water distribution is just if I purchase n litres of water, then it is also just if I purchase n + 1 litres of water.
So (3) The water distribution is just if I purchase all the water.

Since the argument is valid by mathematical induction, and Nozickeans accept the first premiss and deny the conclusion, they must reject the induction step. But to reject the induction step is to accept is negation. The negation of the induction step implies that there was a last just distribution and a first unjust distribution which only differed by the purchase of 1 litre of water. According to principle N, this purchase must have been an unjust step. But how could this step be unjust if there is no significant difference between it and its just predecessor?

Those who concede that this example establishes Quest's contention that N enmeshes Nozick in a sorites paradox might nevertheless deny that the falsity of N immediately follows. For one might point out that the acceptability of N will depend on the stance one adopts to sorites paradoxes in general. However, almost all proposed solutions to the sorites are incompatible with Nozick's position. Those who take sorites vulnerability as disclosing an underlying incoherence in the predicate in question will not have much sympathy with Nozick's position. Nor can Nozick expect a hearty reception from those who believe that classical logic only applies to the Platonic heavens or those who believe it can only be applied to the ordinary world in a rough and ready way. For Nozick's entitlement theory is intended as an analysis of a *coherent* concept, defended by means of logical arguments that hold for the ordinary world. Nor can Nozick be in sympathy with those who hope to solve sorites arguments by rejecting mathematical induction or restricting it to precise contexts. For principle N is itself an inductive principle. Nor will Nozick's position be attractive to those who believe that sorites difficulties show that we must accept truth value gaps. Principle N would extend the determinacy of the early judgements 'Distribution 1 is just', 'Distribution 2 is just', . . . to later statements which are the best candidates for truth value gaps. Since N disallows the relevant sort of truth value gaps, Nozick's position should be unattractive to those who adopt supervaluational logic in response to the sorites. Nozick's theory should also be unappealing to the many-valued theorist. According to the many-valued approach, the induction step's seductiveness is due to the fact that the corresponding conditionals obtained by specifying values for n always have a very high *degree of truth*. However, these tiny departures from complete truth accumulate resulting in the falsity of the induction step. If the mathematical induction were recast as a giant chain of *modus ponens* subarguments, each step would be 'nearly valid'. Each step is almost completely truth preserving but there is some leakage. However, the whole spirit of principle N is to prevent subtle leaks in justice judgements.

> Whatever arises from a just situation by just steps is itself just. The means of change specified by the principle of justice in transfer preserve justice. As correct rules of inference are truth-preserving, and any conclusion deduced via repeated application of such rules from only true premisses is itself true, so the means of transition from one situation to another specified by the principle of justice in transfer are justice-preserving, and any situation actually arising from repeated transitions in accordance with the principle from a just situation is itself just. (p. 151).

The analogy Nozick draws between justice and validity is an analogy between justice and *classical* validity.

The only approach compatible with Nozick's position is the epistemic one. According to epistemic approach, there is an unknowable counterexample to the induction step. There was an unjust 1 litre purchase but it must pass undetected because of the semantic uncertainty induced by vague predicates. Because this crucial transaction was an unjust step, principle N is left unrefuted. The general price of adopting the epistemic approach is the acceptance of an extraordinarily sharp distinction between assertibility and truth conditions along with consequent puzzles about language learning. The only philosophers who have paid this price in print are James Cargile, Richmond Campbell, and Israel Scheffler. The particular price of Nozick's adoption of the epistemic approach (besides the small audience problem) is first felt in the mystery of how we could ever learn the meaning of 'justice'. For according to the epistemic approach, 'just' can apply to a situation without the possibility of us knowing that it applies. How can the meaning of such a word be taught? A second cost peculiar to Nozick's use of the epistemic approach is that the practical utility of N will be compromised. For once we permit unjust steps to be indiscernibly different from just steps, uncertainties about whether N applies will accumulate as the number of prior applications grows. Thus N will not function as an inductive principle normally does in logic and mathematics. The role of N in Nozick's entitlement theory of justice is to make this historical theory plausible by linking together stages of history. Under the epistemic approach to the sorities, N delivers strong links at the metaphysical level but not at the epistemological level. So even according to the entitlement theory, knowledge of all transactions could leave us in ignorance about the justice of a situation because of the semantic uncertainty of 'justice'.

University of Delaware,
Newark, Delaware 19716, U.S.A.

©Roy A. Sorensen 1986

HARRIS ON NON-VICIOUS CHOICE

By Oliver Leaman

JOHN Harris argues that there is a principle of non-vicious choice which permits a parent to save her child rather than a complete stranger (*The Value of Life*, Routledge, 1985, pp. 71–2). The case he has in mind is that of a doctor with a variety of patients to treat who decides to save her own children first. He argues that the doctor's choice to favour her children is not a vicious choice, but is instead benign by contrast, for example, with a preference for her own race. If she has a preference for her own race then she may value the life of a patient of a different race unequally, and he might be the victim of vicious discrimination.

Harris gives two reasons for calling the doctor's preferential treatment of her children benign. The first reason is that this preference generalised would not lead to any group being systematically and unfairly disadvantaged. The second reason is that the disposition to love one's family and friends is on the whole well worth having, resulting in a better and happier world. Neither of these reasons is compelling. The first reason does not recognize that those with few friends and surviving relatives could well constitute a disadvantaged group in any system where benefits are contingent upon being part of a network of such friends and relatives. Harris's first reason would justify an extended system of nepotism, and even those with an extensive set of friends and relatives might on occasion suffer unfair discrimination because no friend or relative was on hand to choose them in preference to others. Harris's second reason requires our acceptance of an unstated move from 'F loves G' to 'F is justified in choosing to benefit G'. We may indeed want to live in a world in which people love their families and friends, and yet not want to live in a world in which they choose to benefit their family and friends when acting in a professional capacity such as that of a doctor. We might well expect a professional person to take steps to ensure that she does not discriminate in favour of the people she loves when she is acting professionally. Presumably that is one of the reasons why applicants for particular types of employment are asked to reveal whether they are related to anyone on the selection committee, and why academic staff are discouraged from having emotional and sexual relationships with their pupils. Such behaviour discriminates, and discriminates viciously, against those applicants and students who are not either related to or loved by those in power and authority over them. If there is a principle of non-vicious choice, Harris has not yet properly established it.

Liverpool Polytechnic,
Barkhill Road,
Liverpool L17 6BD

© OLIVER LEAMAN 1986

PASCAL'S WAGER AND INFINITE UTILITIES

By ANTONY DUFF

PASCAL'S Wager purports to offer us reason to take steps to bring it about that we believe in God, just so long as the probability that God exists is greater than zero. The basic thought is this: if God exists, the reward for the believer is of infinite value (and the penalty for the unbeliever is of infinite disvalue); whereas if God does not exist the believer's loss (if any) and the unbeliever's gain are both finite. Two features of the argument based on this thought concern me here.

First, since we are not certain that God exists, the Wager must be based on the 'expected value' of believing in God. To determine that expected value we would – as good Bayesians – need to multiply the value of believing in God, if He exists, by the probability that He does exist: but for the purpose of the Wager we do not need to assign any particular probability to God's existence – all we need say is that it is greater than zero. Since the value of believing in God is, if He exists, infinite; and since an infinite value multiplied by however small a probability generates an infinite expected value: the expected value of believing in God is infinite if His existence is to the smallest degree probable. The expected value of not believing in God, however, remains finite, however large the probability that He does not exist: so long as the probability that God exists is greater than zero the expected value of believing in God is therefore greater than that of not believing in Him.

Second, the Wager does not give us reason to 'believe in God'; for we cannot simply *decide* to believe in God. We can however take steps which will make it more likely, even highly probable, that we will come to believe in God; we can expose ourselves to influences which are likely to induce belief. This is what the Wager supposedly gives us reason to do; to try to increase the probability that we will come to believe in God; for since believing in God has a greater expected value that not believing, the attempt to induce that belief in oneself has a greater expected value that the alternative of not trying. We cannot admittedly guarantee that we will come to believe: thus to find the expected value of the attempt to come to believe we must multiply the value of believing in God if God exists not just by the probability that He does exist, but also by the probability that we will through this attempt come to believe in Him. But the expected value of this course of action remains infinite – and thus greater than that of the alternative of not trying: for an infinite value generates an infinite expected value however many, and however small, the probabilities by which it is multiplied; infinity times anything greater than zero remains infinity.

There are many objections to this argument. Ian Hacking has argued, however, that the argument is valid; that the conclusion does follow from the premises – though those premises are at best dubious. But in terms of its own logic the argument actually has *no* implications for what we should do; and certainly not the implication that we should take steps to bring it about that we believe in God. For suppose I take no steps to make it more likely that I will come to believe in God. There must still be some probability, however small, that I will nonetheless come to believe in Him – there must be such a probability even if I try to *avoid* coming to believe in Him; and that probability is enough to generate an infinite expected value for my actions. No course of action can make it absolutely certain that I will not come to believe in God: therefore, every course of action has an infinite expected value – the infinite value of believing in God multiplied by the probability that God exists, and by the probability that I will come to believe in Him. I have, therefore, no reason to try to increase the probability that I will come to believe in God, since no such increase in probability can increase the expected value of my actions – which is already infinite. This is indeed a manifestation of God's benevolence: that however I live, whatever I do, the expected value of my actions is infinite! (It might be said that I at least have reason not to kill myself, since my death would certainly prevent my coming to believe in God: but even this is not so. For I cannot guarantee absolutely that anything I do will bring about my death so instantaneously that I have no time to come to believe in God;

there will always be some probability, however small, that my action will not cause my death, or that I will come to believe in God even as I breathe my last breath.)

'But surely, if one of the courses of action available to me makes it more probable that I will gain some hoped-for benefit, the hope of that benefit must give me reason to undertake that course of action. I have a 1/1,000,000 change of gaining £1,000,000 if I do nothing; by speaking one sentence I can increase that chance to 1/10,000 – and all it costs me is the effort of speaking that sentence: obviously I have reason to speak that sentence.'

Of course this is usually true, just because we are usually dealing with finite values: with any finite value an action which increases the probability of gaining that value has (to that extent) greater expected value than one which does not. But it is not true in this case, just because we are dealing with an infinite value: an infinite value confers an infinite expected value on any and every course of action which involves any probability, however small, of gaining that value; increasing the probability of gaining it does not increase the expected value of my action.

This might look like a *reductio ad absurdum* of the Wager. It might more particularly suggest that there is something wrong with trying to capture the infinite in a calculus of probabilities.[1]

The University of Stirling,
Stirling FK9 4LA

© R. A. DUFF 1986

---

[1] This note was provoked by hearing George Schlesinger defend Pascal's Wager in a paper called 'You Bet Your Life': a version of that paper, by George Schlesinger and William Lycan, is due to be published in *Man and Responsibility*, edited by Joel Feinberg. For the logic of Pascal's argument, I have also drawn on I. Hacking, 'The Logic of Pascal's Wager' *American Philosophical Quarterly* 9, 1972, p. 186; and M. Martin, 'On Four Critiques of Pascal's Wager' *Sophia* 14, 1975, p. 1. Thanks to Andrew Brennan and to Murray MacBeath.

REPEATABLE MIRACLES?

By ANDREW REIN

A MIRACLE – by which I mean, following Hume, a violation of a law of nature – is a logical but not a physical possibility. It is a logical possibility because the following scenario is coherent: a deity created the universe and set it in motion according to fixed rules (laws of nature) some of which from time to time he sees fit to break. It is not a physical possibility because its occurrence contravenes one or more physical laws (laws of nature).

Many philosophers who accept the logical possibility of a miracle occurring insist, however, that there could not be a *repeatable* miracle (see, for example, Richard Swinburne, *The Concept of Miracle*, Macmillan 1970, and Ninian Smart, *Philosophers and Religious Truth*, SCM Press 1964, especially pp. 26–7 and p. 37 respectively). I will suggest that there is no interesting sense in which this claim is true.

What claim is being made here depends, of course, on what is meant by 'a repeatable event'. Ordinarily, we say that an event has been repeated when an event of the same kind has occurred.[1] So an event is not repeatable, according to this meaning, if an event of the same kind could not recur. Now it is not in this ordinary sense that the claim is usually intended but it is instructive at the outset to see why, on this understanding of the claim, it is false. It is false because the conception which guarantees the logical possibility of a miracle occurring also guarantees the logical possibility of its recurrence. If a deity sees fit on some occasion to break a natural law in bringing about a certain event then it must be within his power to break that natural law on another occasion in order to bring about an event of the same kind. (Of course, some events, by their very nature – for example, the running of the first four minute mile – cannot be repeated. But I am disregarding such events here.) So, for example, it must be within God's power to effect a parting of the Red Sea today if it suits His purpose even though (we assume) He effected a similar miracle years ago. Of course, the more often a deity affects miracles of the same kind, the more chaos he creates and the less secure we will feel about the natural law or laws he has chosen to flout. But this does not affect the coherence of the claim that the natural laws are in operation and that the events in the natural world would conform to them if only the deity would stop interfering.

If the claim that miracles are unrepeatable is not to be understood in this ordinary way (according to which it would thus seem plainly false) how then is it to be understood? Let us take our cue from Swinburne (*The Concept of Miracle* pp. 26–7) who writes:

> There cannot be repeatable counter-instances to genuine laws of nature, that is, counter-instances which would be repeated in similar circumstances. Repeatable counter-instances to purported laws only show those purported laws not to be genuine laws.

Swinburne's point here can be better appreciated if we gloss his notion of 'repeatable' as follows. Suppose that according to one or more natural laws an event of a certain kind E is expected to occur in and only in a given set of circumstances. Now suppose that in these circumstances an event s of kind S occurs which is incompatible with the occurrence of an *E*-type event. Since *s* is an individual event there will be some feature of its occurrence which marks it off from all previous *E*-type events. An event *s'* is a repeat of *s* if and only if *s'* is an *S*-type event the occurrence of which is marked by that same feature which marked *s* off from all previous *E*-type events.

---

[1] What are the criteria for one event being of the same kind as another? The question probably contains a mistake but in any case I do not need to answer it here.

Given this account, the claim that a miracle cannot be repeated has a compelling defence, as follows. The natural laws which predicted the occurrence of an *E*-type event when in fact *s* occurred were simply inaccurate because they failed to account for the fact that in certain special circumstances (namely, those characterising the occurrences of *S*-type events) *S*-type events rather than *E*-type events occur. If, indeed, we succeed in isolating these special circumstances, then we will be in a position to modify our natural laws so that they correctly predict when *E*-type events will occur and when *S*-type events will occur. Thus any repeated violation of a purported natural law is not a violation of a genuine law (i.e., not a miracle) but only shows that the purported law requires modification.

The difficulty with this argument is that it makes it appear as though any set whatsoever of special circumstances which characterise the occurrences of a series of apparently repeated miracles will count against their authenticity. But this cannot be right – as I will try to illustrate by means of the following example. Suppose a man has been observed on a number of occasions restoring the sight of men who, according to current ophthalmic knowledge, are incurably blind. (We may assume that their optic nerves have perished.) Let us call this worker of apparent miracles 'Jesus'. Jesus' method is this: he lightly touches the blind men's eyelids with his fingertips, he mutters a few words, and on opening their eyes the men find that they can see again. Other people attempt this method but meet with no success.

Now it is easy to see that the apparent miracles all occur in circumstances which are similar in one crucial respect: Jesus is present. Thus, assuming that Jesus is never present when the miracle is attempted but fails, we have here a hypothetical situation in which an apparent miracle has been repeated (according to the account of this notion under consideration). But these apparent miracles are surely no less authentic for that.

In the hypothetical situation just outlined, not only are there special circumstances which single out these apparent miracles but we know what they are; so we can frame a rule which will enable us successfully to predict when another such apparent miracle will occur. Such a rule, however, would never acquire the status of a scientific explanation of the phenomena it predicts. This is because it is so totally at odds with the prevailing laws of ophthalmology that we are left utterly mystified as to the processes involved in the restoration of sight to the incurably blind. Until we are provided with some insight into these processes we would continue to regard the healing of the blind as genuinely miraculous despite our success in predicting when another such miracle will occur. And the reason for this, I suggest, is just that the rule by means of which we make our predictions falls so far short of providing a scientific explanation of the phenomena it predicts – that is, falls so far short of being accepted as a law of nature.

The upshot of this then is that it is not just any set of special circumstances marking off a series of apparently miraculous events which tells against their authenticity. It is only when that set of special circumstances generates a regularity embodied in a natural law that their authenticity is undermined.

Smart, I believe, is aware of this point, for it appears to motivate his claim that miracles are not *experimentally* repeatable (See *Philosophers and Religious Truth* p. 37). If a series of controlled experiments has produced the same anomalous result then, according to Smart, this amounts to 'a small-scale law of nature: it is described by something of the form "Under these special circumstances, that occurs"' (p. 48). An if that is so then the result of the experiments is not anomalous after all but merely indicates that some revision is required in some large-scale natural law or laws. Now it must be agreed that we probably would conclude from the result of our experiments that a small-scale natural law is governing them. But we need not do so. Suppose that a deity has interfered with our experiments. He has not done this according to any fixed plan or rule; it is just that each time an experiment has been conducted he decided on the spur of the moment to intervene. This is a whimsical deity indeed but not an impossible one. And the possibility of such a deity establishes the possibility of experimentally repeatable miracles.

To conclude. In the senses of 'repeatable' with which we have been concerned, a repeatable miracle could (logically) occur. If, however, we mean by a repeated event that its recurrences are governed by a regularity enshrined in a natural law then there can be no repeatable miracles. But this claim is trivial. For if the occurrence of an event is governed by the regularity enshrined in some natural law, then that occurrence cannot be in violation of the laws of nature (assuming, of course, these laws to be consistent) and hence cannot be a miracle.[2]

University College,
Oxford OX1 4BH

© ANDREW REIN 1986

[2] I wish to thank Paul Ziff for helpful comments.

CORRECTION

The following footnote was inadvertently omitted from Jonathan Harrison's article 'Distribution, superposition and quantum logic' in ANALYSIS 45.4 October 1985, pp. 204-7: 'Thanks for help are due to my colleague Robert Black, who, however, does not share the views expressed in the article.'

CALL FOR PAPERS

The editors of **Social Philosophy & Policy** are pleased to announce that Volume 6, Issues 1 and 2 will be open to unsolicited manuscripts.

Issue 1 will be devoted to the theme of "Capitalism" and Issue 2 to the theme of "Socialism."

For information concerning deadlines for submission and a style sheet, write to:

John Ahrens, Managing Editor
SOCIAL PHILOSOPHY AND POLICY
Social Philosophy and Policy Center
Bowling Green State University
Bowling Green, OH 43403
USA

Basil Blackwell • Oxford and New York

PHILOSOPHICAL BOOKS

Edited by Anthony Ellis

Philosophical Books was founded by the **Analysis** committee in 1960 to provide prompt, scholarly reviews to assist both librarians and individuals in the choice of professional works of philosophy. Besides conventional reviews in both contemporary philosophy and the history of the subject, the journal includes regular discussion features where authors of selected titles have the opportunity to reply to reviewers. There are also occasional subject surveys, and issues devoted to a single theme.

January, April, July, October
Subscription Rates Volume 27, 1986
Individuals: £13.15 (UK) £14.45 (overseas) US$28.00 (N America)
Institutions: £27.50 (UK) £34.35 (overseas) US$55.50 (N America)

Please send orders with payment to:

Iris Taylor
Journals Department
Basil Blackwell
108 Cowley Road
Oxford OX4 1JF
England

or to: Journals Department
Basil Blackwell
PO Box 1320
Murray Hill Station
NY 10156
USA

Basil Blackwell • Oxford and New York

PHILOSOPHICAL INVESTIGATIONS

Edited by D Z Phillips

Philosophical Investigations features articles, reviews, and in depth critical notices by experts such as Cora Diamond, Frank Ebersole, Antony Flew, Raimond Gaita, Oswald Hanfling, Lars Hertzberg and Rush Rhees.

January, April, July, October
Subscription Rates Volume 9, 1986
Individuals: £14.25 (UK) £17.85 (overseas) US$30.35 (N America)
Institutions: £32.50 (UK) £40.65 (overseas) US$65.55 (N America)

Please send orders with payment to:

Iris Taylor
Journals Department
Basil Blackwell
108 Cowley Road
Oxford OX4 1JF
England

or to: Journals Department
Basil Blackwell
PO Box 1320
Murray Hill Station
NY 10156
USA

Basil Blackwell • Oxford and New York

---

forthcoming

PHILOSOPHICAL TOPICS

Volume 14, Number 1 Spring 1986
Topic: **Epistemology**
Issue Editor: **Roberta Cutler Klein**

Contributors: **Jonathan Adler, William Alston, Robert Audi, Laurence Bonjour, Roderick Chisholm, Peter Klein, Keith Lehrer, Ernest Sosa, Peter Unger, and Michael Williams**

Volume 14, Number 2 Fall 1986
Topic: **Value Theory**
Issue Editors: **Jacob Adler and Richard N. Lee**

Contributors: **Hugo Bedau, Stephen Darwall, John Martin Fischer, Alan Gewirth, R.M. Hare, Thomas E. Hill Jr., Jeffrie Murphy, Peter Railton, Michael Slote, and Susan Wolf.**

Philosophical Topics appears twice yearly. Each issue is devoted to questions in one of the major areas of philosophy and consists entirely of solicited papers.

Subscription Rates: Individual $20, Institutional $30. All checks should be made payable to The University of Arkansas Foundation, Inc./*Philosophical Topics* and sent to the Managing Editor at the Department of Philosophy, 127 Kimpel Hall, University of Arkansas, Fayetteville, Arkansas 72701.

NOTES

The ANALYSIS Committee consists of: Chairman, P. T. Geach; Secretary, T. R. Baldwin; T. S. Champlin, M. Clark, J. P. Dancy, Andrew Harrison, B. J. Heal, R. F. Holland, J. H. McDowell, D. H. Mellor, H. W. Noonan, P. Smith, M. A. Stewart, R. G. Swinburne, A. R. White, C. J. F. Williams, Peter Winch. This committee is responsible for appointing and advising the Editor and for the general policy of the paper.

SUBSCRIPTIONS. ANALYSIS is published by Basil Blackwell Ltd, 108 Cowley Road, Oxford OX4 1JF, England and P.O. Box 1320, Murray Hill Station, NY 10156, U.S.A. Each volume comprises four numbers, each of at least 48 pages, appearing in January, March, June and October. Orders should be sent to the publisher. Subscriptions prices for 1986 are as follows:

	U.K.	Overseas (incl. Eire)	U.S.A. & Canada
Institutions	£11.75	£14.50	U.S. $33.00
Individuals	£7.85	£9.75	U.S. $21.00

BACK ISSUES. Volumes 1–37 are now available in complete volumes in an offset reprint from Swetz & Zeitlinger, 471 Keizersgracht, Amsterdam, Holland. Recent back issues are available from the publisher at the current single issue price.

MICROFILM copies of complete volumes are available from Volume 1 onwards from University Microfilms International Ltd, 30–32 Mortimer Street, London W1N 7RA, England, or 300 North Zeeb Road, Ann Arbor, Michigan 48106, U.S.A.

ADVERTISING. For details contact Bill Drake, telephone (0869) 38477, or write c/o the publishers.

CONTRIBUTIONS. Articles submitted for publication should be addressed to Mr Christopher Kirwan, Exeter College, Oxford OX1 3DP.

Contributors are asked to note the following requirements. Articles should normally not exceed 3,000 words in length, including references and footnotes. They should be type written in double spacing on one side of the paper only; photographic copies are acceptable. Footnotes should be kept to a minimum and wherever possible avoided altogether. Inset quotations and formulae should be well spaced from the text preceding and following them. The spacing of symbols within formulae should be carefully attended to. Single quotation marks should be used, except for purposes of internal quotation and "scare" quotes. Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

Typescripts of rejected articles will be returned only if self-addressed envelopes with correct postage are provided (in the case of overseas contributors, international reply coupons sufficient for airmail).

Proofs of accepted articles will be sent to authors for correction, together with information about free copies. Typescripts of accepted articles will be retained by the Editor on the assumption that authors have kept their own copies.

Apart from correction of printers' errors, only very minor changes in proof will normally be accepted by the Editor, to avoid delay and expense in production.

Typeset by Unicus Graphics Ltd, Horsham, Printed and bound in Great Britain by
Whitstable Litho Ltd, Whitstable

Vol. 46 No. 3 (New Series No. 211) June 198

ANALYSIS

Edited by

CHRISTOPHER KIRWAN

CONTENTS

The contingent *a priori*: has it anything to do with indexicals? TIMOTHY WILLIAMSON
Epimenides and Curry LAURENCE GOLDSTEIN
On the logical possibility of time without change S. G. WILLIAMS
Redundancy and *INUS* conditionality THEODORE C. DENISE
Causality and the supposed counterfactual conditional in Hume's *Enquiry* ANNE JAAP JACOBSON
Why Chisholm's analysis of justification won't do C. KENNETH WATERS
Kvanvig and Swain on the basing relation LORY LEMKE
Acting intentionally and minimal abilities MICHAEL J. COSTA
Prosthesis and pre-emption JAMES MONTMARQUET
Ways of speaking of imagination ALAN R. WHITE
Metaphysical realism and historical interpretation PETER SMITH
Noonan on naming and predicating E. J. LOWE
The avoidance of the problem of evil: a reply to McGrath ROGER CRISP

BASIL BLACKWELL PUBLISHER · OXFORD

ISSN 0003-2638

THE PHILOSOPHICAL QUARTERLY

Executive Editor: John Hall

Founded by the Scots Philosophical Club and the University of St Andrews in 1950, **The Philosophical Quarterly** aims to foster and publish significant contributions in every branch of the subject, promoting discussion of recent philosophical works, and providing critical reviews of recent books, in a style accessible to as wide a group of philosophers as possible.

The April 1986 issue is a Special Issue entitled **Mind, Causation and Action**, and includes the following articles:

Invited introduction: finding psychology
SIMON BLACKBURN

Causes of behaviour
MICHAEL MORRIS

Mental causes and the explanation of action
GRAHAM and CYNTHIA MACDONALD

The failure of Lewis's functionalism
JOSEPH OWENS

Wittgenstein on sensuous experience
MALCOLM BUDD

Knowing what we think
WILLIAM CHARLTON

Incontinent believing
ALFRED R. MELE

Russell's deceptive desires
GEORGE GRAHAM

Deliberation and the presumption of open alternatives
TOMIS KAPITAN

Published January, April, July, October for the Scots Philosophical Club and the University of St Andrews
Subscription Rates Volume 36, 1986
Individuals: £11 (UK) £13.75 (overseas) US$26 (N America)
Institutions: £27.50 (UK) £34.35 (overseas) US$57.50 (N America)

Please send orders with payment to:

Iris Taylor
Journals Department
Basil Blackwell
108 Cowley Road
Oxford OX4 1JF
England

or to: Journals Department
Basil Blackwell
PO Box 1320
Murray Hill Station
NY 10156
USA

Payment can be made by cheque or money order made payable to Basil Blackwell Ltd, by credit card (Access, American Express, Barclaycard, Diners Club, Mastercharge, Visa; please state card number and expiry date), or through the National Girobank (account number 236 6053). Sample copies, single and back issues are available; please write for details.

Basil Blackwell · Oxford and New York

METAPHILOSOPHY

Edited by Terrell Ward Bynum

Metaphilosophy publishes articles and book reviews in a wide range of philosophical topics including the foundations, scope and function of philosophy, aspects of philosophical systems, applied philosophy, the sociology of philosophy, the teaching of philosophy, presuppositions of philosophical schools, and the relation of philosophy to other disciplines.

Volume 16 Number 4 was a special issue on Computer and Ethics, containing the following articles:

What is computer ethics?
JAMES H. MOOR

Should computer programs be private property?
DEBORAH G. JOHNSON

Responsibility for computer-based error
JOHN W. SNAPPER

Attributing responsibility to computer systems
WILLIAM BECHTEL

Frankenstein's children: artificial intelligence and human value
DAN LLOYD

January, April, July, October
(two single issues and one double issue per year)
Subscription Rates Volume 17, 1986
Individuals: £19.80 (UK) £24.75 (overseas) US$41.80 (N America)
Institutions: £36 (UK) £45 (overseas) US$75 (N America)

Special offer for new subscribers – we will send you a copy of Computers and Ethics completely free when you subscribe to Volume 17!

Please send orders with payment to:

Iris Taylor
Journals Department
Basil Blackwell
108 Cowley Road
Oxford OX4 1JF
England

or to: Journals Department
Basil Blackwell
PO Box 1320
Murray Hill Station
NY 10156
USA

Payment can be made by cheque or money order made payable to Basil Blackwell Ltd, by credit card (Access, American Express, Barclaycard, Diners Club, Mastercharge, Visa; please state card number and expiry date), or through the National Girobank (account number 236 6053). Sample copies, single and back issues are available; please write for details.

Basil Blackwell • Oxford and New York

ANALYSIS 46.3 JUNE 1986

THE CONTINGENT *A PRIORI*: HAS IT ANYTHING TO DO WITH INDEXICALS?

By TIMOTHY WILLIAMSON

I

'CAN some contingent truths be known *a priori*?': when this question is raised in modern philosophy – as, following Kripke, it often has been – it generally introduces a discussion of certain examples which seem to turn on indexical or indexical-like words (Nathan Salmon lists most of the important work on the topic at pp. 77–8 of *Reference and Essence*, Blackwell 1982; I have also been helped by Graeme Forbes's typescript 'States of Affairs, *Bedeutung* and the Contingent *A Priori*', 1985, and subsequent correspondence). Sometimes the indexicality is quite obvious, as in 'I am here now', sometimes it appears only on analysis, as in 'If anyone uniquely invented the zip, Julius did', where by stipulation 'Julius' rigidly designates the inventor of the zip (if there is one), for the rigidity of 'Julius' can be expressed by means of the 'actually' operator, which is at least formally analogous to an indexical (cf. Martin Davies *Meaning, Quantification, Necessity*, Routledge 1981, pp. 230–8 on the similar metre rod example). Moreover, this indexicality has been exploited in a well known argument that what is known *a priori* can be contingent only in a superficial way (Gareth Evans 'Reference and Contingency' *Monist* 62 (1979) pp. 161–89; technical details are omitted, since deep and superficial contingency coincide for sentences, such as those considered below, without 'actually' or similar operators). The purpose of this note is to argue that there are deeply contingent *a priori* truths which in no way turn on indexicals. I do not deny that there are also contingent *a priori* truths which *do* turn on indexicals, but we are unlikely to get clear on the contingent *a priori* in general if we take it to be essentially tied up with indexicals when in fact it is not.

Consider the following example:

(1) There is at least one believer

To keep (1) free of indexicals, the verb in it should be read tenselessly. A believer is anyone who has beliefs. For the moment, I assume (1) to be a contingent truth: I show later how it can be modified to meet someone who thinks that at least one believer – God – necessarily exists. Now since it is impossible to believe (1) falsely, it seems that one can know (1) *a priori*. Note that (1) is contingent in the deep sense in which Evans suggests that a contingent *a priori* truth *would* be persistently puzzling (technically, possible worlds semantics can evaluate (1) as false even with respect to a world which it is treating as actual): if there is a paradox about the contingent *a priori*, it has not been dissolved by Evans.

However, it might be replied that, although (1) does not contain indexicals, one can know (1) *a priori* only if one infers it from the vaguely Cartesian proposition:

(2) I am a believer

Therefore, it might be continued, one's knowledge of (1) is *a priori* only if one's knowledge of (2) is *a priori*. Moreover, since (2) entails (1), (1) is contingent only if (2) is contingent. Thus, the objection would go, (1) is contingent *a priori* only if (2) is contingent *a priori*, and (2) does involve an indexical. At this point, the objection could be rounded off either by a mild moral or by a severe one. The *mild moral* would concede that (2) is a contingent *a priori* truth, but claim that this turns on its indexicality, that I could not know the truth of what results from replacing 'I am' in (2) by 'N.N. is' (where 'N.N.' is my name); thus the link between the contingent *a priori* and indexicals would remain. Note that this mild moral would not help Evans, for (1) and (2) remain deeply contingent *a priori* truths in his sense. The *severe moral* (which is inconsistent with the mild one) would deny that (2) can be known *a priori* at all, and consequently that (1) can be known *a priori* either, on the grounds that (2) can be known only by *a posteriori* introspection, observation of one's own behaviour or the like. This severe moral would help Evans, since neither (1) nor (2), however deeply contingent, would be *a priori*.

I shall answer both lines of objection by denying that one can know (1) *a priori* only by inferring it from (2). Consider the following method for forming beliefs:

(M) Given a valid deduction from the premise that someone believes that *P* to the conclusion that *P*, believe that *P*

It is logically necessary (and, therefore, *a priori*) that any belief which one forms by use of (M) is true. Since (M) is an absolutely reliable method for forming true beliefs, any reasonable theory of knowledge should count a belief that *P* which has been formed by use of (M) as a case of knowledge that *P*. Moreover, since one can know *a priori* that (M) is absolutely reliable, the knowledge which it gives should be *a priori* (note also that (M) appeals only to ratiocination). In other words, if one comes to believe that *P* by use of (M), one knows *a priori* that *P*.

Now replace '*P*' in (M) by (1). In order to apply (M) here, one needs a valid deduction to the conclusion (1) from the premise:

(3) Someone believes that there is at least one believer

Such an inference is, of course, trivially valid. Now the crucial point is that (M) does *not* require us to derive *a priori* knowledge of (1) from *a priori* knowledge of (3). If it did, we should be in difficulties. For *a priori* knowledge of (3) might in turn have to be derived from *a priori* knowledge of:

(4) I believe that there is at least one believer

But the attempt to derive *a priori* knowledge of (1) from *a priori* knowledge of (4) (via (3)) is open to just the same objections as were levelled against the attempt to derive it directly from *a priori* knowledge of (2); in both cases, the indexical 'I' would be crucial. Fortunately, (M) avoids these problems. For it gives the valid deduction from (3) to (1) as *by itself* a sufficient reason to believe (1); (4) is irrelevant. Thus one can come to believe (1) by use of (M): so, by the conclusion of the previous paragraph, one can know the contingent truth (1) *a priori*. This way of coming to know (1) did not exploit indexical or indexical-like words at any point.

(1) can be modified to accommodate a religious believer. Consider:

(1′) There is at least one fallible believer

One is fallible, in the required sense, if and only if not all of one's beliefs are true. Presumably, it is possible that no believer other than God exists and impossible that God is fallible: hence (1′) is a contingent truth. In order to use (M) to gain *a priori* knowledge of (1′), one needs a valid deduction to the conclusion (1′) from the premise:

(3′) Someone believes that there is at least one fallible believer

Here is such a deduction: suppose that (3′) is true; consider any *x* who believes that there is at least one fallible believer; either *x* is fallible or *x* is not fallible; if *x* is fallible, *x* is a fallible believer, so (1′) follows by existential introduction; if *x* is not fallible, all of *x*'s beliefs are true, so *x*'s belief in (1′) is true; either way, (1′) is true. Thus one can come to believe (1′) by use of (M): so, by the conclusion of the previous paragraph but one, one can know the contingent truth (1′) *a priori*. This way of coming to know (1′) did not exploit indexical or indexical-like words at any point. Incidentally, the reader might care to imagine the following situation: there are (up to now) no believers except for a necessarily not fallible one, who is now in the act of coming to know (1′) *a priori*.

Intuitively, (1) is a much *less* puzzling case of the contingent *a priori* than, for example, 'If anyone uniquely invented the zip, Julius did'. With the latter, a linguistic stipulation seems to summon substantive knowledge out of thin air, and Evans's deflationary tactics are apposite. In contrast, (1) is the kind of proposition which we should *expect*, for good Kantian reasons, to be contingent *a priori*; those reasons are articulated in (M).

II

In arguing that (M) gives *a priori* knowledge, I appealed to our grasp of the concept of the *a priori*, but not to any formal definition. This was because it is very hard to find a non-trivial definition of the *a priori* which does not clearly deviate in extension from the concept to be defined. Indeed, traditional definitions seem not to apply to any *necessary* truths, let alone contingent ones. For example, Kant himself, in the *Critique of Pure Reason*, defines knowledge to be *a priori* if it is 'absolutely independent of all experience' (B3, Kemp Smith's translation). Is *our* mathematical knowledge absolutely independent of such experiences as checking through a proof we had written out some time before, or listening to a teacher tell us that our answer to an exercise was correct? (cf. Paul Teller 'Computer Proof' *Journal of Philosophy* 77, 1980, pp. 797–803; Michael Detlefsen and Mark Luker 'The Four Color Theorem and Mathematical Proof' *ibid.* pp. 803–20; Philip Kitcher *The Nature of Mathematical Knowledge*, (Oxford University Press 1983)). Tamara Horowitz ('A Priori Truth' *Journal of Philosophy* 82, 1985, pp. 225–39) has recently given a new definition, on the basis of which she accepts examples of the contingent *a priori* similar to (1) (*ibid.* p. 233). Roughly, Horowitz defines a truth *p* to be *a priori* if and only if everyone necessarily has enough experience to know that *p* whenever they exist and have the concepts in *p* and *p* is true. However, this condition seems to be neither sufficient nor necessary for a truth to be *a priori*. For let *p* be the proposition that no one has the concept of a joke; since anyone who has the concepts in *p* has the concept of a joke, the truth of *p* is incompatible with anyone's having the concepts in *p;* thus Horowitz's conditional is vacuously satisfied, since its antecedent is self-contradictory; hence had *p* been true, as it could have been, it would have met the condition – but it would not have been *a priori*, for it is impossible to know that *p*. Conversely, someone could have the concepts in a mathematical truth *p* without in Horowitz's sense having enough experience to know that *p*, because they did not have all the concepts in any *proof* of *p*, even though such proofs existed; hence *p* would fail the condition – but it would still be *a priori* (for Horowitz, mathematical truths are paradigms of the *a priori*). Similar examples to the latter also suggest that Philip Kitcher's definition (op. cit. p. 24) does not give a necessary condition for *a priori* knowledge, although it may give a sufficient one (if so, it could be used to defend the claim that (M) yields *a priori* knowledge – in Kitcher's terminology, the use of (M) could be an *a priori* warrant). I certainly know of no definition of the *a priori* which is plausible enough to be relied on as a *test* of whether a given truth is *a priori*.

However, to claim that (M) guarantees *a priori* knowledge is in effect to lay down the following condition as *sufficient* for *P* to be knowable *a priori*: it is impossible to believe falsely that *P* (I assume bivalence). One then naturally wonders whether this condition is not also *necessary*.

Trinity College,
Dublin 2

©TIMOTHY WILLIAMSON 1986

EPIMENIDES AND CURRY

By LAURENCE GOLDSTEIN

MOST modern discussions of the Liar paradox focus on a version, or a version of a version, attributed to Eubulides, e.g. 'This statement is false'. The version attributed to Epimenides the Cretan – 'All Cretans are liars' – is largely ignored and, at first sight, this disregard seems warranted. For Epimenides' statement seems not paradoxical, but merely *false*. If we take Epimenides to be saying

(E1) All statements made by Cretans are not true

then (E1), itself a statement by a Cretan, cannot be true; yet it can be false just in case some Cretan made a true statement. However, as Alonzo Church pointed out [2], the Epimenides has a paradoxicality that the Eubulides lacks: since (E1) *must* be false, it follows that some true statement *must* have been made by some Cretan. And it is surprising that the *empirical* question of whether any statement other than (E1) was made by a Cretan is settleable by pure logic.

A more dramatic example is given by L. J. Cohen [3]. Imagine a court hearing in which a policeman testifies that anything which the prisoner deposes is false. Then the prisoner takes the stand and deposes that something which the policeman testifies is true. It follows, as Cohen shows, that something which the policeman testifies is false, and something which the prisoner deposes is true. And from *this* it follows, as Arthur Prior noticed, that either the policeman or the prisoner must have said something *besides* the assertions we reported them as making. Prior remarks '... these *two* assertions cannot have even been made – not just cannot have been true – unless at least *three* statements have been made by these people. Yet it does not seem *logically* impossible that the policeman and the prisoner should make the assertions mentioned and no others (the court house having been bombed immediately afterwards, say).' ([9], pp. 71-72, see also [10] and [11]).

What is peculiar about the Epimenides and its variants is that they seem to show that some or other empirical circumstance (that makes an 'additional' proposition true) *must* obtain, and yet we know independently that such circumstances *need not* obtain. Whereas, in the Eubulidean paradoxes, the paradox-reasoning leads us to a contradiction, in the Epimenidean paradoxes, logical reasoning leads to a conclusion which is not contradictory but which, on empirical grounds, we might wish to reject as false.

* * *

It is safe to assume that, if there was an ancient Cretan who said that all Cretans are liars, he meant to exempt himself, or at least that very statement, from the generalization. Yet from the statement 'All statements made by Cretans, excluding this one, are not true', no paradox ensues. For paradoxicality, the self-referential case needs to be included, and we can do this explicitly by rewriting (E1) as

> (E2) This statement is not true, and all other statements made by me and other Cretans are not true.

This formulation illustrates the relation between the Epimenidean and Eubulidean paradoxes (see [7]). (E2) is not Eubulides-paradoxical if there is *any* true Cretan statement, for then (E2) is not true, and, from the non-truth of (E2) we may *not* infer its truth.

It is quite clear that the reference to utterances of *Cretans* in (E2) is not essential to the paradox, reference to *any* proposition will preserve paradoxicality. Let 'q' be *any* proposition. Then

> (E3) This statement is not true and q

is Epimenides-paradoxical, for (E3) cannot be true (otherwise it would be true that it is not true, hence it would, *per impossible*, be both true and not true) so it *must* be not true, and hence '$\sim$ q' is necessarily true. The argument seems to show that a proposition *must* have a certain truth value which we know independently it need not have – and this is the hallmark of Epimenides-paradoxicality.

(E3) contains all that is essential to an Epimenides paradox, and we can write out a slightly more formal version of the paradoxical sentence as follows:

E: E is not true and q.

As we remarked, what is distinctive about such paradoxes – what makes them paradoxical – is that, given only that an Epimenidean statement has a truth value, it seems that we can prove an arbitrary proposition true.

* * *

The Epimenides paradox, when seen in this light, is strongly reminiscent of Curry's paradox [4]. It can be shown that consideration of the class $\{x: x \in x \rightarrow q\}$ devastates any attempt to construct a non-trivial naive set theory in which the full comprehension principle holds, for such a theory can be proved *absolutely* inconsistent (any arbitary proposition can be derived within it) using only a very weak logic [8]. Semantical versions of Curry's paradox appear equally devastating. Consider the simple case (for a more elaborate case, see [5]) of the statement

If this statement is true, then q

where 'q' is an arbitrary proposition and the 'if ... then' is a material conditional. The above statement cannot be *false*, for a material conditional is false only when its consequent is false and its antecedent is true. And the antecedent says that the statement is *true*. So the conditional must be *true*. Hence its antecedent is true, so its consequent, 'q', must be true too. Again, we appear to have a logical proof that some arbitrary proposition is *true*.

When we write out the formal version of the Curry statement (using the usual definition '$\sim A \vee B$' for the material conditional), the relationship between the Curry and the Epimenides paradoxes is indeed striking. We have

C: C is not true v q (Curry)

E: E is not true & q (Epimenides).

Here also is the Liar in a 'strengthened' Eubulidean version

L: L is not true

I discussed L in another number of this journal [6]. What I argued there was that one cannot simply 'let' a statement be one that says that that statement is not true, for that would be to *identify* a statement with a statement having the opposite truth-value. Now, in the case of the Curry statement C, one can let C be the statement 'C is not true v q' only if both C and 'q' are true, for otherwise the two will have different truth values. In the case of the Epimenides E, one can let E be the statement 'E is not true & q' only if both E and 'q' are false. In both cases, therefore, the paradox-reasoning only shows that the propositions must have the truth-values that are necessary for setting up the examples.

But just how do we fail to 'set up the example' if, say, in an Epimenidean case, we substitute for the arbitrary proposition a proposition that is *true*? What I have claimed is that we cannot 'let' a statement S be

S is not true and snow is white.

The reason is that the truth value of the conjunction of some proposition with a truth is the truth value of that proposition, so to let S be the above statement is tantamount to stipulating that S have the same truth value as its own negation – and we are simply not empowered to make such stipulations. No statement can be such that it and its negation have the same truth value, so 'S' designates no statement. It might be thought strange that whether something is a statement or not depends on the colour of snow. Yet clearly *there is no statement* K such that

K is not true and K is true

so there is no statement M such that

M is not true and M has the same truth value as 'Snow is white'.

It is an empirical fact which determines that there is no statement fitting the above description, and *calling* something a statement (e.g. 'This statement is not true and has the same truth value as "Snow is white" ') doesn't *make* it one.

* * *

We can now see why *self-referentiality* (of a direct or indirect kind) invariably features in semantical paradoxes. In the vast literature on the subject, nobody (to my knowledge) has asked *why* semantical paradoxes are (directly or indirectly) self-referential, although many writers, noting the presence of self-reference in the paradoxes, have sought to ban it or to find some rule to prevent it arising (e.g. Russell's Vicious Circle Principle – see [1], chapter 1).

Our argument has been that paradoxes arise when a putative statement is said or implied to be both true and not true, and that paradox is dissipated by first discovering that this *is* what is said or implied and, second, noting that there can *be* no statement that satisfies this pair of ascriptions. Now, in a statement about the truth value of *another* statement, the truth value of the statement doing the referring will depend on the truth value of the statement to which reference is made. For example, if statement X says that statement Y is not true, then, assuming that both X and Y are statements, they must have opposite truth values. But in a self-referential case, where the statement doing the referring is the same as the statement being referred to, we may get truth and non-truth ascribed to *one and the same statement*. Consider the Epimenidean example

E: E is not true and q.

Here 'q' is *both* conjoined with a statement about E, and also *part of* E (so, although 'q' is arbitrary, it is not independent of E). If 'q' is not true, then, in virtue of E *containing* 'q' as a conjunct, E is not true, and this is consistent with what the first conjunct (the statement *about* E) says. However, if 'q' is true then we have the absurd (and thus to be rejected) implication that E is both true and not true.

Hong Kong University,
Pokfulam Road,
Hong Kong

© LAURENCE GOLDSTEIN 1986

REFERENCES

[1] C. Chihara, *Ontology and the Vicious Circle Principle*, Ithaca: Cornell University Press, 1973.
[2] A. Church, 'Review of A. Koyré's *The Liar*', *Journal of Symbolic Logic*, 12 (1946), p. 131.
[3] L. J. Cohen, 'Can the logic of indirect discourse be formalized?', *Journal of Symbolic Logic*, 22 (1957), pp. 225-32.
[4] H. B. Curry and R. Feys, *Combinatory Logic*, Amsterdam: North Holland Publishing Co., 1958.
[5] P. T. Geach, 'On Insolubilia', *Analysis* 15.3, January 1955, reprinted in his *Logic Matters*, Oxford: Basil Blackwell, 1972, pp. 209-11.
[6] L. Goldstein, 'The paradox of the liar – a case of mistaken identity', *Analysis* 45.1, January 1985, pp. 9-13.
[7] J. L. Mackie, 'Conditionally Restricted Operations', *Notre Dame Journal of Formal Logic*, 2 (1961), pp. 236-43.
[8] R. K. Meyer, R. Routley and J. M. Dunn, 'Curry's Paradox', *Analysis* 39.3, June 1979, pp. 124-8.
[9] A. N. Prior, 'Epimenides the Cretan', *Journal of Symbolic Logic*, 23 (1958), reprinted in his *Papers in Logic and Ethics*, London: Duckworth, 1976, pp. 70-7.
[10] A. N Prior, 'On a family of paradoxes', *Notre Dame Journal of Formal Logic*, 2 (1961), pp. 16-32.
[11] A. N. Prior, 'A Budget of Paradoxes', in his *Objects of Thought*, Oxford: Oxford University Press, 1971.

ON THE LOGICAL POSSIBILITY OF TIME WITHOUT CHANGE

By S. G. Williams

I

THE question whether there can be time without change does not seem on the face of it to be any easier to answer than other substantive questions concerning the nature of time. But if Arthur Prior, and following him, John Lucas are to be believed, it can in fact be answered using a simple technical argument (see J. R. Lucas, *A Treatise on Time and Space*, Methuen 1973, pp. 10–11; and A. N. Prior, *Papers on Time and Tense*, Oxford 1968, p. 111). In this paper, I wish to consider this argument in detail. I shall try to show that the argument does not establish its conclusion, and that its failure is almost certainly the product of equivocation; I shall conclude with a brief model theoretic conjecture as to what may have prompted the equivocation. The version of the argument which I shall consider will essentially follow Lucas's, though unlike his it will be concerned with the possibility of time after the end of change rather than time before its beginning. (In this respect I follow Prior). Nothing of importance hangs on the difference.

II

In the form in which I want to consider it, the argument seeks initially to show that the thesis that time will end is not entailed by the thesis that change will end, at least in a certain minimal tense logic. By exploiting the method used in establishing this, it then seeks to draw the conclusion first that the two theses are logically distinct, and secondly that time without change is logically possible.

In more detail the argument runs as follows. Let 'Fp' mean 'it will sometime be the case that p', and 'Gp' mean 'it will always be the case that p'. Then in the context of an unrestricted rule of substitution, the thesis that time will come to an end is plausibly represented by

$$(1)\ [Fp \rightarrow FGq];$$

while the thesis that change will come to an end is plausibly represented by

$$(2)\ [Fp \rightarrow F[q \rightarrow Gq]].$$

(These representations are not unassailable; but for the purposes of the present note, I shall assume that they are adequate.)

Now (1) entails (2) even in Lemmon's minimal tense logic, K_t. But the converse does not hold. For letting 'P' and 'H' be past tense operators corresponding to 'H' and 'G', K_t consists of a standard propositional calculus base, two new rules of inference, $\vdash \alpha \Rightarrow \vdash G\alpha$ and $\vdash \alpha \Rightarrow \vdash H\alpha$, and four new axioms:

$$[G[p \rightarrow q] \rightarrow [Gp \rightarrow Gq]] \qquad [H[p \rightarrow q] \rightarrow [Hp \rightarrow Hq]]$$

$$[PGp \rightarrow p] \qquad [FHp \rightarrow p].$$

(See e.g. Prior, op. cit., p. 99.) So if we interpret all of the tense operators as 'unary null operators' (see Lucas op. cit., p. 11), all the rules of K_t preserve tautologousness, all its axioms become tautologies, (2) becomes:

(3) $[p \rightarrow [q \rightarrow q]]$,

which is also a tautology, and (1) becomes:

(4) $[p \rightarrow q]$,

which is not. So (2) cannot entail (1) in K_t.

But then (so the argument goes) this implies that (1) and (2) are logically distinct: for it follows that upon some interpretation of the propositional variables and tense operators, the axioms of K_t and (2) are true even though (1) is false. And since (1) is false upon an interpretation if and only if its negation is true, it follows also that upon exactly the same interpretation, the axioms of K_t, (2) and the negation of (1) are all true. Hence time without change is logically possible.

III

This argument – which I shall refer to as the *Prior-Lucas argument* – is (I submit) fallacious. The first part of it does of course succeed in showing that (2) does not entail (1) in K_t. This follows from the proven fact that all the theorems of the theory consisting of K_t and (2) have a property – tautologousness under the relevant interpretation of the tense operators – which (1) does not. Moreover, it is a consequence of this that there is some interpretation of the propositional variables and tense operators according to which all the axioms of K_t and (2) are true, and (1) is false; and that upon the same interpretation, all of the axioms of K_t, (2) and the negation of (1) are true.

Without further argument, however, it does not follow that (1) and (2) are logically distinct, nor does it follow that time without change is logically possible. For although it is possible to interpret the propositional variables and the tense operators in K_t, (2) and (1) so that all the axioms of K_t and (2) are true and (1) is false, and although the reinterpreted version of the negation of (1) would then be true, these reinterpreted sentences cannot be used to represent the original temporal theses – indeed they are not essentially about temporal matters at all. Hence although (3) may be logically distinct from (4), and the conjunction of (3) and the negation of (4) may be logically possible, we cannot (for all the argument shows) draw the same conclusions about (1) and (2) on the one hand, and the conjunction of (2) with the negation of (1) on the other. (1) and (2) mean entirely different things from (4) and (3). (It would be like trying to argue that 'The Pope is a bachelor' and 'The Pope is an unmarried man' are logically distinct, and that 'Some bachelors are married' is logically possible, on the grounds that the first two sentences would have different truth values, and the third would come out true, if 'bachelor' were reinterpreted as 'married man'. And this of course is preposterous.)

Of course I do not wish to deny that there may be a sense of logical distinctness and logical possibility according to which it is indeed sufficient for proving the distinctness of (1) and (2) and the possibility of time without change that one prove the model theoretic results above. But should someone employ the concepts of logical distinctness and logical possibility in just this way, he must not be misled into thinking that he has shown that (1) does not entail (2) *per se*, or that for all that can be known *a priori* time without change is possible. And in the present context, these are surely the only conclusions of interest.

IV

The above considerations show that the Prior-Lucas argument fails; but they suggest in addition a natural explanation of the failure, namely that the argument equivocates. For it depends (so it seems) on treating (1) as the same thesis as (4), (2) as the same thesis as (3), and the characteristically tense logical fragment of K_t as that same fragment but with the tense operators reinterpreted. And each of these assumptions is false.

But now this is such an obvious mistake. How could it possibly have arisen? I shall conclude with a brief suggestion as to how it might have done so, a suggestion involving a subtle and easily made confusion in model theory.

Let $S_1, \ldots, S_k$, S be sentences in a language L, and let $e_1, \ldots, e_n$ be expressions of L. Then let us say that the set $\{S_1, \ldots, S_k\}$ is *interpretationally consistent* relative to the meanings of $e_1, \ldots, e_n$ just in case there is a possible admissible interpretation of L which assigns to $e_1, \ldots, e_n$ their intended interpretations and upon which $S_1, \ldots, S_k$ are all true; and let us say that the sequent $S_1, \ldots, S_k \vdash S$ *admits of an interpretational counterexample* relative to the meanings of $e_1, \ldots, e_n$ just in case there is a possible admissible interpretation of L which assigns to $e_1, \ldots, e_n$ their intended interpretations and upon which $S_1, \ldots, S_k$ are true and S is false. (Notice that the constant expressions $e_1, \ldots, e_n$ need not come from the so-called logical constants; they can be expressions of any sort.)

Then using this terminology, the Prior-Lucas argument shows first that the sequent,

(5) $AX(K_t), (2) \vdash (1)$

(where '$AX(K_t)$' represents the axioms of K_t) admits of an interpretational counterexample relative to the meaning of '→'; and secondly that the set,

(6) $\{AX(K_t), (2), neg(1)\}$

(where 'neg' represents a sentential negation operator) is interpretationally consistent relative to the meanings of '→' and 'neg'. And as we have seen already, this does not establish that (1) and (2) are logically distinct or that time without change is logically possible. In order to do *those* things (in the present context) the argument would have to show that (5) admits of an interpretational counterexample relative to the meanings of '→', 'F', 'P', 'G' and 'H', and that (6) is interpretationally consistent relative to the meanings of '→', 'neg', 'F', 'P', 'G' and 'H'. And if what I have said so far is right, it certainly does not do that.

However, while the argument may not establish these latter claims, it is nevertheless very easy to think that it does, on the one hand by mixing up the expressions relative to whose meanings the sequent (5) admits of a counterexample, and on the other by mixing up the expressions relative to whose meanings the set (6) is consistent. And mixing up the expressions in this way generates precisely the equivocation upon which the Prior-Lucas argument almost certainly depends. My suggestion, then, as to what may have prompted the equivocation is that the first pair of model theoretic assertions has been mistaken for the second as a result of just such a mixing up process. Mistakes like these are not unsubtle, not uncommon, and very easy to make.

V

The Prior-Lucas argument, I submit, fails; and it fails (so it seems) because it equivocates. I have tried to isolate the natural source of that equivocation, and this may or may not be right. But whatever the explanation of the equivocation, it seems certain that the question of whether there can be time without change is no easier to answer (for all that the argument genuinely establishes) than other substantive questions concerning the nature of time.[1]

Oriel College,
Oxford OX1 4EW

© S. G. Williams 1986

---

[1] I am grateful to Richard Swinburne, David Charles and the Editor for comments on earlier drafts of this note.

REDUNDANCY AND *INUS* CONDITIONALITY

By THEODORE C. DENISE

ON reading A. J. Dale's '*INUS* Conditions' (ANALYSIS 44.4, October 1984, pp. 186–8), a response to my 'On the Nature of *INUS* Conditionality' (ANALYSIS 44.2, March 1984, pp. 49–52), it becomes clear that my proposed amendment to the established definition of an INUS condition (J. L. Mackie, 'Causes and Conditions', *American Philosophical Quarterly* 2, October 1965, p. 246) needs clarification and additional support. In that paper, I both acknowledge that Frank Jackson shows the established definition, Mackie's definition, to be too broad ('Reviews', *Journal of Symbolic Logic*, 47, June 1982, p. 472) and claim that my amendment narrows it without corrupting its basic structure. The amended definition stands to the original as a precising definition.

By way of orientation, I shall state briefly four items which are very much at point. The first and second: the established definition, with my amendment inserted between dashes, is this:

> A is an *INUS* condition of a result P if and only if, for some X and for some Y, $(AX$ or $Y)$ is a necessary and sufficient condition of P – and A is a necessary condition of PX – but A is not a sufficient condition of P and X is not a sufficient condition of P.

The third: Mackie's sole explanation of his definitional provision against redundancy – his *minimum sufficient conditions* proviso – is this: '[If] ... "$AB\bar{C}$" represents a sufficient condition of the fire, and one that contains no redundant factors ... [then] ... $AB\bar{C}$ is a minimal sufficient condition for the fire' (op. cit., p. 246). The fourth: Jackson's device for exposing the inadequacy of the established definition is set up thusly: With $S_1, S_2, \ldots, S_n$ marking distinct sufficient conditions of P and their disjunction constituting a necessary and sufficient condition of P, i.e., $(S_1 \vee S_2 \vee \ldots \vee S_n) \equiv P$, fix $A(X)$ as $A(S_1 \vee \bar{A})$ and Y as $(S_2 \vee S_3 \vee \ldots \vee \bar{A}S_1)$.

Section I explains that I do survive Dale's attempted *coup de grâce*; section II directs us back to the redundancy problem of INUS-theory; section III reasserts, more accountably, the need for my PX-only-if-A amendment; and the appendix removes Jackson's device from disrepute.

I

According to Dale, to employ Jackson's device is to neglect Mackie's minimum sufficient conditions proviso: any demonstration of an example running counter to the established INUS definition would contain, say, $\bar{A}S_1$; but this disjunct 'is not a minimum sufficient condition since S_1 is *ex hypothesi* sufficient alone.' It would seem therefore that any counterexample produced through this device must pertain solely to the established definition shorn of its minimum sufficient conditions proviso.

The quick, defensive response to Dale's fell swoop argument is this: Jacksonian demonstrations have near analogues to which he doesn't object in demonstrations by Jaegwon Kim ('Causes and Events: Mackie on Causation', *Journal of Philosophy*, 68, July 1971, pp. 426–41). Let us observe this for $\bar{A}S_1$ with S_1 as BC where we are bound by Kim's most immediate measure for Mackie's understated proviso, *viz.*,

> The event $A_1A_2 \ldots A_n$ is a minimal sufficient condition for an event P if and only if it is a sufficient condition for P, and, for each $i(1 \leqslant i \leqslant n)$, $A_1 \cdots A_{i-1} \cdot A_{i+1} \cdots A_n$ is not a sufficient condition for P (ibid., p. 433).

The Jacksonian move BC to $\bar{A}BC$ is prohibited, but the Kimian move to the $\bar{A}BC$ equivalent $\bar{A}(A \vee B)C$ is not. (Jackson's device is for locating within a set of logically equivalent statements those with a certain characteristic; it is *not* for constructing statements. See Appendix.)

Confident that my amendment is in answer to a definitional problem eliminated from the start, Dale attempts to show that it is wrong-headed as well. He argues both that it grants INUS status to irrelevant conditions and that it withholds INUS status from relevant ones. The first argument requires us to agree along the way that A is shown to be a necessary condition of PX because of the logical fact that a contradictory PX 'entails A (or any other sentence for that matter!).' Confusions aside, this argument is without impact. The term 'necessary condition' is being used in a truth-functional sense which simply cannot serve as the image for any eliminative induction sense of the term.

Dale begins the second argument by drawing attention to Mackie's early comments about the 'at least an INUS condition' concept and he ends by attributing absurdity to my modified definition in connection with a switching circuit problem. There are things to object to throughout. Thus, for example, he is beginning with a concept that is presumptive of the established definition and is ending with a problem that is anything but innocent (see Appendix). The hinge of the argument is, however, clear enough: if in saying that A causes B one intends the sense that A is (at least) an INUS condition of P, then, by the modified definition, he is 'implicitly asserting that $\bar{A}$ is a necessary condition for each of the other minimum sufficient conditions.' Dale is claiming, for example, that if there are three disjuncts in $(AX \vee Y) \equiv P$ and A is a necessary condition of PX, then $(AX \vee \bar{A}X' \vee \bar{A}X'') \equiv P$. But this is not so: if $\bar{A}X'$, then X and X' represent distinct complexes of conditions, i.e., $X \neq X'$, since A is a necessary condition of PX. Now assume $\bar{A}X'$; this leaves (1) $X \neq X' = X''$, (2) $X = X'' \neq X'$, and (3) $X \neq X' \neq X''$. If $\bar{A}X''$, then, if (1), assumed $\bar{A}X'$ is reiterated, and if (2), A's being a necessary condition of PX is contradicted (3) does not preclude $\bar{A}X''$, of course, but neither does it preclude AX''.

II

Mackie's functional account of the minimum sufficient conditions proviso comes to this: a statement of the necessary and sufficient conditions of P in a field F of candidate conditions must be in disjunctive normal form (*DNF*) with respect to these conditions (op. cit., p. 255); furthermore, the thusly involved disjuncts of conditions must be redundancy-free as determined by standards of eliminative induction (ibid., pp. 246–7 and 255–7). The root problem here is that the first part is about the truth-functional logic of the language expressing complexes while the second part is about assessing them. As Kim points out, however, it would seem that a solution is gained by advancing a step beyond Mackie (op. cit., pp. 434–5): the second part becomes commensurate with the first if we insist that an INUS-cogent statement is not only in DNF with respect to conditions but also that *its DNF portion is simpler than any DNF equivalent.*

Kim has second thoughts about this version of the proviso even as he suggests it. He is aware – as was Mackie, no doubt – of a 'peculiarity' in the logic of normal forms. Consider: (i) with favorable observations $(AB\bar{C} \vee A\bar{B}C \vee \bar{A}B\bar{C}) \equiv P$, we determine by logical methods that, with respect to DNF portions, $(B\bar{C} \vee A\bar{B}C) \equiv P$ is *the* simplest DNF, but (ii) with favorable observations $(AB\bar{C} \vee A\bar{B}C \vee A\bar{B}\bar{C} \vee \bar{A}BC \vee \bar{A}B\bar{C}) \equiv P$, we determine comparably that, with respect to DNF portions, both and only $(A\bar{B} \vee \bar{A}B \vee A\bar{C}) \equiv P$ and $(A\bar{B} \vee \bar{A}B \vee B\bar{C}) \equiv P$ are the simplest DNFs. By the suggested version of the minimum sufficient conditions proviso, (i) provides an INUS-cogent statement but (ii) does not. Beyond counselling caution about this version because its analysis is wanting, Kim reacts to its preclusion of (ii) as weighing against it. This is strange. Why should we think that 'yes' is the answer to the following question? If it is not anomalous that neither of the initial statements of (i) and (ii) are INUS-cogent, is it anomalous that no statement logically equivalent to the latter is INUS-cogent even though, admittedly, one statement logically equivalent to the former is?

III

My amendment is in response to an unremarked type of redundancy, the *redundancy of severally determined kinds of P-occurrences*. Two of the facts accounting for it are these: first, although DNF statements are disjunctive in the inclusion sense, they may or may not have mutually exclusive component disjuncts. Second, there are sets of logically equivalent DNF statements wherein only some members have mutually exclusive component disjuncts.

Stringent though the unique simplest DNF version of the minimum sufficient conditions proviso may seem, it is not stringent enough to (fully) protect the established definition of an INUS condition. What is required beyond this – in line with Mackie's efforts – is, I believe, my amendment. Consider the following two arrays of regular occurrences to which I have added, for each, a monitoring feature for its kinds of P-occurrences as determined by the conditions of F:

			Case I		Case II	
A	B	C		P		P
p	p	p		a	$\{a\}$	p
p	p	a	$\{a\}$	p	$\{b\}$	p
p	a	p	$\{b\}$	p	$\{c\}$	p
p	a	a		a		a
a	p	p		a	$\{d\}$	p
a	p	a	$\{c\}$	p		a
a	a	p		a	$\{e\}$	p
a	a	a		a		a

(K is the set of P-occurrences. For I: $K = \{a, b, c\}$ and the assigned denotations of conditions are $A = \{a, b\}$, $\bar{A} = \{c\}$, $B = \{a, c\}$, $\bar{B} = \{b\}$, $C = \{b\}$, $\bar{C} = \{a, c\}$. For II: $K = \{a, b, c, d, e\}$, $A = \{a, b, c\}$, $\bar{A} = \{d, e\}$, $B = \{a, b, d\}$, $\bar{B} = \{c, e\}$, $C = \{a, c, d, e\}$, $\bar{C} = \{b\}$.) With 'p' for 'present' and 'a' for 'absent' and with the unit subsets of K located by row, each array provides for making an initial statement with the conditions in expanded DNF and for constructing a partition (π) of K. Let us refer to such a statement as the *primal* DNF (I am widening 'DNF' to cover biconditionals with their conditions portion in DNF), and to such a partition as the *finest* π. We identify a set of logically equivalent DNFs by means of the former and a family of various refinement by means of the latter.

Case I. The primal DNF is $(AB\bar{C} \vee A\bar{B}C \vee \bar{A}B\bar{C}) \equiv P$ and the finest π is $\{\{a\}, \{b\}, \{c\}\}$. *The* simplest DNF is $(B\bar{C} \vee A\bar{B}C) \equiv P$; it provides for constructing the somewhat refined π $\{\{a, c\}, \{b\}\}$.

Case II. The primal DNF is $(ABC \vee AB\bar{C} \vee A\bar{B}C \vee \bar{A}BC \vee \bar{A}\bar{B}C) \equiv P$ and the finest π is $\{\{a\}, \{b\}, \{c\}, \{d\}, \{e\}\}$. *The* simplest DNF is $(AB \vee C) \equiv P$; it does *not* provide for constructing a π, but it does for constructing $\{\{a, b\}, \{a, c, d, e\}\}$.

Since the simplest DNFs of I and II are unique, they describe circumstances satisfying the established definition replete with the stringent version of the minimum sufficient conditions proviso. Moreover, since the simplest DNF of I provides for constructing a π of K, it describes a circumstance satisfying the amended definition as well (the disjuncts are mutually exclusive). The test imposed by the amendment amounts to this: for all A and X determinations on the basis of the simplest DNF, it is the case that X *does not occur without* A *in the primal DNF.*

In contrast, since the simplest DNF of II does not provide for constructing a π of K, it describes a circumstance *not* satisfying the amended definition. The simplest DNF $((AB \vee C) \equiv P)$ is not free from the redundancy of severally determined kinds of P-occurrences (the disjuncts are not mutually exclusive). Its illegitimacy as an INUS-cogent statement is suggested by an item of thought it supports: kind of P-occurrence $\{b\}$ is both distinguishable and not distinguishable from kind $\{a\}$ with respect to their minimum sufficient conditions. (Notice that some nonunique simplest DNF situations, e.g., the one in Section II, involve instances of this redundancy.)

APPENDIX

Jackson's device serves singularly in identifying instances of the redundancy of severally determined kinds of P-occurrences. Consider our Case I: from INUS-cogent candidate $(B\bar{C} \vee A\bar{B}C) \equiv P$ we move to $(AB\bar{C} \vee A\bar{B}C \vee \bar{A}B\bar{C}) \equiv P$ by the device. We then note that neither $(AB\bar{C} \vee A\bar{B}C) \equiv P$ nor $(A\bar{B}C \vee \bar{A}B\bar{C}) \equiv P$ follows and judge that, as so far tested, there is no instance of the redundancy. Consider our Case II in contrast: from INUS-cogent candidate $(AB \vee C) \equiv P$ we move to $(ABC \vee C \vee AB\bar{C}) \equiv P$. We then note that, while $(ABC \vee C) \equiv P$ does not follow, $(C \vee AB\bar{C}) \equiv P$ does, and judge that there is an instance of the redundancy.

Let me confirm my remark that Dale's switching problem is too complicated for an absurdity argument. Dale's set-up: 'a switching circuit composed of switches A$'$ and B$'$ in series and these in parallel to C$'$ and D' in series.' ('A is the condition that A$'$ is closed etc.') Fact: $(AB \vee CD) \equiv P$ is *the* simplest DNF (the disjuncts are not mutually exclusive). Comment: from this INUS-cogent candidate we move by the device to $(AB(\bar{C} \vee \bar{D}) \vee CD \vee AB(\bar{C} \vee \bar{D})) \equiv P$. We then note (rewriting) that $(AB\bar{C} \vee AB\bar{D} \vee CD) \equiv P$ follows even though $(CD \vee ABCD) \equiv P$ does not, and judge that there is an instance of the redundancy of severally determined kinds of P-occurrences.

The last word belongs to Mackie: '[In the case of] ... the man who is shot by a firing squad, and two bullets go through his heart simultaneously ... the requirements of my analysis are not met: each proposed cause *is* redundant' (op. cit., p. 251).

541 Hall of Languages,
Syracuse University,
Syracuse, N.Y. 13210, U.S.A.

© THEODORE C. DENISE 1986

CAUSALITY AND THE SUPPOSED COUNTERFACTUAL CONDITIONAL IN HUME'S *ENQUIRY*

By ANNE JAAP JACOBSON

HAVING told us that we may define a cause to be '*an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second,*' Hume disarmingly goes on to say, 'Or, in other words *where, if the first object had not been, the second never had existed,*' [3], p. 76. No correct account can have it that the 'in other words' are *just* other words. Since the added remark relates absences of causes to absences of effects, it must at least, when joined to the first definition, tell us that the causal relationship given in the first definition is really symmetrical (cf. [4, p. 173]). According to a traditional account, Hume's comment plays a rather different role. On this interpretation, the 'other words' give us an analysis of causality which, because it licenses counterfactuals, is said to be incompatible with the regularity account of causality taken to be the main account of causality Hume gives us ([2], [6]).

Making a break with this tradition, Martin Bunzl has argued that while Hume's account of causality is a regularity analysis, that does not mean the use of counterfactuals is unavailable to it [1]. My break with the tradition is more extensive. On the correct account of Hume's thought, I have argued, Hume does *not* give us a regularity account of causality and that there is *no* reason to suppose that the account he does give us disbars him from a use of counterfactual conditionals [5]. However, this interpretation does not make the attribution of the 'other words', when read counterfactually, unproblematic for, as we will see, the 'other words' do not give us a counterfactual conditional.[1] Once we see the correct interpretation of the first definition, it becomes clear that the 'other words' can be read as giving us a factual statement which presents as additional information only the idea that the relation described in the first definition is symmetrical.

It is common to think that the first definition tells us that events like the cause, when they occur, are *always* – throughout all space and time – conjoined with events like the effect ([2], [9]). It is also common to think that this is what Hume means in saying that cause and effect are constantly conjoined. However, there are good reasons for regarding these common constructions as quite wrong. When Hume introduces the relation of constant conjunction in the *Treatise* he says

---

[1] I am indebted to Martha Bolton's querying of the standard reading of Hume's peculiar other words. J. L. Mackie [7] alludes briefly to 'a certain interpretation' (which he appears to reject) of Hume's remark and the context suggests he there has in mind a factual, not counterfactual, reading.

> ... this new-discover'd relation of a constant conjunction seems to advance us but very little in our way. For it implies no more than this, that like objects *have always been* plac'd in like relations of contiguity and succession ... ([4], p. 88; my stress).

Constant conjunction is introduced by Hume as constant conjunction *so far* and this restriction is explicit or clearly implicit in nearly every occurrence of 'constant conjunction' in those sections of the *Treatise* devoted to discussing causality (for a more lengthy discussion of this point than that below see [5]). That this is what we should expect is clear once we consider how Hume is using 'constant conjunction'. While Hume holds that experience itself never informs us of the unobserved – a point crucial to his discussion – he says that experience can inform us of a constant conjunction. Furthermore, Hume holds that statements of constant conjunction are deductively insufficient to give us statements about the unobserved, a point inexplicable on the standard reading. For example, Hume says

> ... but even after experience has inform'd us of their *constant conjunction*, 'tis impossible for us to satisfy ourselves by our reason, why we shou'd extend that experience beyond those particular instances, which have fallen under our observation [4, p. 91].

> Thus tho' causation be a *philosophical* relation, as implying contiguity, succession, and constant conjection, yet 'tis only so far as it is a *natural* relation, and produces an union among our ideas, that we are able to reason upon it or draw any inference from it [4, p. 94].

But the first definition in the *Treatise* is said to give us a philosophical relation and so it, and its near duplicate in the *Enquiry*, should *not* be read as enabling us to draw inferences about what has yet to occur. That is, we should expect the first definition to say no more than that objects like the cause have so far been conjoined with objects like the effect. And Hume explicitly says that this is the correct reading of the first definition in the *Enquiry*. In the *Enquiry* Hume tells us what the first definition means when applied to the statement 'This vibration is the cause of this particular sound' and what it means, we are told, is '... that this vibration is followed by this sound, and that all similar vibrations *have been* followed by similar sounds' ([3, p. 77]; my stress).

Once we see that the first definition of causality is about the past and remind ourselves that 'had' is the pluperfect and may be read simply as signifying completion in the past [8], we can find an easy 'factual' reading for the supposedly counterfactual conditional. So read, the conditional is different from 'where, if the first *has* not been, the second never has existed' in its indication that the relevant series is completed. Compare: a teacher remarks concerning a student just given a viva, 'Had Catherine ever had formal instruction in logic? I don't know if she ever in fact received such instruction, but I do know that if she had not, she had not had any need for it'. To say this is not at all to commit oneself to what would counterfactually have happened had she not received instruction when she in fact had had it. Similarly, I remark about a friend, 'Had he visited London while he was in England? I don't know if he had, but I do know that if he had not, that was not because he hadn't wanted to.' Here again, the conditional is not counterfactual. Equally similarly, Hume need not be read as concerned with what would have happened had a cause occurred in circumstances in which it did not. Grammar is not decisive since we can read '... had ... had ... ' as counterfactual; for example, 'the effect had occurred had the cause not failed to occur.' But once the temporal restrictedness of the first definition is noticed, the factual reading of the added comment becomes very attractive. And it allows us an appealing answer to the question, 'What besides symmetry is brought in by this untypical, never repeated sentence?' The answer is: nothing.

Rutgers University, Davison Hall,
Douglass Campus, New Brunswick,
NJ 08903, U.S.A.

© ANNE JAAP JACOBSON 1986

REFERENCES

[1] Martin Bunzl 'Humean Counterfactuals', *Journal of the History of Philosophy*, 1982, 171-7.

[2] Antony Flew *Hume's Philosophy of Belief*, (Routledge and Kegan Paul, 1961), 129-39.

[3] David Hume *The Enquiry Concerning Human Understanding*, 2nd edition, ed., L.A. Selby-Bigge.

[4] David Hume *A treatise of Human Nature*, the Selby-Bigge edition as revised by Nidditch.

[5] Anne Jaap Jacobson 'Does Hume Hold a Regularity Theory of Causality?', *History of Philosophy Quarterly*, 1984, 75-91.

[6] D. Lewis 'Causation', *Journal of Philosophy*, 1973, pp. 556-67; reprinted in Ernest Sosa, ed., *Causation and Conditionals*, Oxford University Press, 1975.

[7] J. L. Mackie *The Cement of the Universe*, Oxford University Press, 1974, p. 30.

[8] *The Compact Edition of the Oxford English Dictionary*, p. 1126 (p. 127 of vol. *H* in the standard edition).

[9] J. A. Robinson 'Hume's Two Definitions of "Cause",' *The Philosophical Quarterly*, 1962; reprinted in V. C. Chappell, ed., *Hume*, Macmillan, 1966, pp. 129-47.

WHY CHISHOLM'S ANALYSIS OF JUSTIFICATION WON'T DO

By C. KENNETH WATERS

RODERICK Chisholm tries to rescue the traditional analysis of knowledge as justified true belief by formulating a new definition of justification (Chisholm 1982, pp. 43-9). His attempt to save the traditional analysis fails, however, because, as I will show, it rests on a definition of justification that rules out the possibility that interesting false beliefs might be justified. Ironically, in the same volume where he presents his new definition of justification, Chisholm notes the possibility of having justified false beliefs. 'Perhaps there was a time,' writes Chisholm, 'when people were justified in believing the false proposition that all swans are white.' (Chisholm 1982, p. 4) I will show that this proposition, like nearly all false propositions, could not possibly be justified according to Chisholm's own definition. Hence, what Chisholm calls justification is not justification at all.

If Chisholm's concept of justification does not capture what we mean by that term, then Chisholm has failed to show that knowledge is justified true belief. This does not necessarily imply that his analysis of knowledge is wrong, for knowledge may not be justified true belief. But it does show that contrary to what Chisholm would like to believe, his concept of knowledge is not closely aligned to the traditional analysis.

Chisholm offers his definition of justification in order to explain why certain propositions in Gettier-type examples are not justified. Gettier-type examples are supposed to provide instances of justified true beliefs that do not constitute knowledge. The examples often involve disjunctions that derive their alleged justification from a false disjunct and their truth from an unjustified disjunct. Chisholm illustrates the Gettier problem with the case of Jones and Brown. In this example, Smith seems to have a justified true belief, namely that Jones owns a Ford or Brown is in Barcelona. This disjunction is called *h*. It turns out that Smith's evidence for *h* relates only to the false disjunct, *f*, 'Jones owns a Ford'. The true disjunct, 'Brown is in Barcelona', is true by coincidence; Smith has no reason to believe it. Smith's evidence for *f* (Jones owns a Ford) is made up of a conjunction of *true* propositions, *e*, which includes conjuncts such as 'Jones keeps a Ford in his garage,' 'Jones says he owns a Ford,' 'Jones has always been honest about such things,' etc. Since, according to this example, Smith is justified in believing the true conjunction *e* and since *e* provides very strong evidence for *f* (Jones owns a Ford), Smith is justified in believing *f*. Of course Smith doesn't know *f* because *f* is false. Since the justified belief *f* implies *h* (Smith owns a Ford or Brown is in Barcelona), *h* must also be justified. So *h* is a justified true belief. Yet, how could we say that

Smith knows *h*? After all, he only believes *h* because he believes its false disjunct.

Chisholm tries to account for the Jones and Ford example by arguing that although Smith is justified in believing the conjunction *e* (Jones keeps a Ford in his garage and Jones says he owns a Ford, etc.), Smith is not justified in believing the disjunction *h* (Jones owns a Ford or Brown is in Barcelona). That is, Chisholm formulates a definition of justification according to which *e* is justified and *h* is not.

On Chisholm's view, a justified proposition must be 'evident.' But being evident is not enough, he suggests, because the proposition might be 'defectively evident.' Roughly speaking, a defectively evident proposition is one that is made evident by a basis which also makes evident a false statement. His formal definition of defectively evident presupposes his concept of basic proposition which he defines as follows:

> *b* is a *basic proposition* for S $=_{\mathrm{df}}$ *b* is evident for S; and everything that makes *b* evident for S entails *b*. (Chisholm 1982, p. 46)

Chisholm defines defectively evident as below:

> *h* is *defectively evident* for S $=_{\mathrm{df}}$ (i) there is a basic proposition for S that makes *h* evident for S and does not logically imply *h*; and (ii) every such basic proposition makes evident a proposition that is false. (Chisholm 1982, p. 46)

Requiring justified beliefs to be nondefectively evident would take care of the Gettier examples because propositions like the disjunction *h* are defectively evident. But as Chisholm explains, this would be requiring too much. For such a requirement would mean that true evident statements such as the conjunction *e*, which is made evident by a proposition which also makes falsehoods such as *h* evident, would not be justified. So he tries to formulate a requirement that will render *e* justified and *h* unjustified.

Chisholm realizes that a conjunction can make evident a falsehood that is not made evident by any one of the individual conjuncts alone. Accordingly, his definition of justification makes a defectively evident conjunction justified if each of its conjuncts is nondefectively evident. Of course, any proposition entailed by such a conjunction is also justified according to his account. Chisholm formulates his definition of justification as follows:

> S is *justified* in believing *p* $=_{\mathrm{df}}$ (i) it is evident for S that *p*; and (ii) if it is defectively evident for S that *p*, then the proposition that *p* is entailed by a conjunction of propositions each of which is evident but not defectively evident for S. (Chisholm 1982, p. 46)

On this definition, the conjunction *e* (i.e. 'Jones keeps a Ford in his garage and Jones has been seen driving a Ford and Jones says he owns a Ford ... '), although defectively evident, is justified (and hence known). The disjunction *h* (i.e. 'Jones owns a Ford or Brown is in Barcelona') apparently is not. The defectively evident conjunction *e* is justified because (i) it is evident; and (ii) it is entailed by a conjunction (e.g. itself) whose conjuncts are each evident but not defectively evident. The conjuncts of *e* are not defectively evident presumably because none of them is made evident by a basic proposition that makes evident a false proposition. In contrast, the defectively evident disjunction *h* appears to fail Chisholm's second requirement for justification because it does not seem to be entailed by any such conjunction.

Chisholm's general strategy for solving the Gettier problem is to redraw the division between justified and unjustified beliefs so that propositions such as 'Jones owns a Ford or Brown is in Barcelona' will no longer count as justified. The problem with his solution, however, is that the division he draws does not separate the appropriate beliefs. For instance, in the Gettier example under consideration, the division should separate *h* (Jones owns a Ford or Brown is in Barcelona) from the well justified (but false) *f* (Jones owns a Ford). The justificatory status of *h*, on his account, however, is linked to the status of *f*. That is, since *f* entails *h*, we can conclude that any conjunction that entails *f* will entail *h*. Actually, given Chisholm's special definition of entailment (see Chisholm 1982, p. 103), this conclusion rests on the following assumption: for any conjunction, if the conjunction is necessarily such that whoever accepts it accepts *f*, then the conjunction is also necessarily such that whoever accepts it accepts *h*. Given the conclusion that any conjunction that entails *f* also entails *h*, *h* must satisfy Chisholm's definition of justification if *f* does. Therefore, since Chisholm claims that *h* is not justified, he is committed to the position that *f* is also not justified. But the conjunction *e* was formulated so that it would provide very strong evidence for *f* (even though *f* turned out to be false). Chisholm manages to explain why Smith does not know that Jones owns a Ford or Brown is in Barcelona only at the cost of holding that Smith is not justified in believing that Jones owns a Ford (no matter how strong the indirect evidence for that belief is).

The failure of Chisholm's definition of justification appears not just with Gettier-type examples, but with any interesting example of a justified false belief. For according to his analysis, practically the only false beliefs that could be justified are ones entailed by basic propositions. This can be proven as follows. Consider any false, evident belief that is not entailed by a basic proposition (e.g. all swans are white). Call it *F*. Since *F* is not entailed by a basic proposition and is false, it must be defectively evident. In addition, since *F* is false, any conjunction that entails it must have at least one false conjunct. Provided that the false conjunct(s) is not implied by a basic proposition, the conjunct, if evident at all, will be defectively evident (since the basis of the false conjunct would make evident the false conjunct itself). Hence, *F* could not satisfy Chisholm's second condition for justification. The only false propositions that could be justified on his account are those that are either entailed by basic propositions or entailed by conjunctions whose false conjuncts all happen to be entailed by basic propositions. Surely we have, or at least have the possibility of having, justified false beliefs that do not meet this stricture. In fact, it seems that any interesting example of a justified false belief, including Chisholm's own example concerning the colour of swans, fails to meet this constraint.

Chisholm's definition of justification is unacceptable because it implies that interesting false beliefs cannot be justified. As mentioned above, this does not mean that his definition of knowledge is necessarily wrong, but it does refute his claim to have 'retained' the traditional definition of knowledge (Chisholm 1982, p. 46). And this might raise doubts, at least among those who worry about finagled conceptions of knowledge whose form relates only to Gettier-type examples. The feat of aligning his definition of knowledge with the traditional one would have lent credibility to the idea that Chisholm was on the right track. Without that close alignment, however, Chisholm's definition appears to be a gerrymandered patchwork of epistemic notions.[1]

John Carroll University,
University Heights, Ohio 44118, U.S.A.

© C. Kenneth Waters 1986

REFERENCE

Chisholm, Roderick M.: 1982, The Foundations of Knowing (University of Minnesota Press, Minneapolis)

---

[1] I would like to thank Bo Clark, Earl Conee, Roderick Chisholm, and John Pollock for helpful comments and criticisms of an earlier version of this paper. I am also grateful for suggestions offered by an anonymous referee and the Editor of *Analysis*.

KVANVIG AND SWAIN ON THE BASING RELATION

By LORY LEMKE

CAUSAL accounts of knowledge are multifarious. A favourite place for inserting causality into an account of knowledge is the basing relation between a belief and its evidence. The case for a conditional causal (or explanatory) account of belief-basing and the well-known gypsy lawyer counterexample to it goes back to 1970 in an exchange between Gilbert Harman and Keith Lehrer in *The Journal of Philosophy* (see volume 67, 1970, pp. 841-55, esp. p. 848, and volume 68, 1971, pp. 311-13). More recently, Marshall Swain has criticized Lehrer's case and affirmed a version of a conditional causal account of belief-basing (see *Reason and Knowledge*, Cornell 1981, chapter 3). And just lately Jonathan Kvanvig has extended Lehrer's type of counterexample against Swain's account (see ANALYSIS, 45.3, June 1985, pp. 153-8).

I will offer a two-part critique of Kvanvig's paper. I do not intend to defend any causal account, but to clarify the issues concerning belief-basing and to indicate what remains to be done for a critic of causal accounts. In the first section, I outline the specifics of a causal account of belief-basing. This is necessary for appreciating my allegation that both Swain and Kvanvig undermine the critical point of Lehrer's counterexample. In the second section, I consider Kvanvig's interesting argument against Swain's conditional causal account. I attempt to show that this argument fails. I conclude that the case against causal accounts of belief-basing is no better off than the position of Keith Lehrer in 1971 – and this is not very satisfying.

I

Let's take the following schema: *e* is conclusive evidence that *i* is true, but *S*'s belief that *e* (*Sbe*) does not cause *S* to come to believe that *i*; whereas *c* is bad evidence for *i*, and *S*'s belief that *c* (*Sbc*) causes *S* to come to believe that *i*. In this case, *S comes to believe* that *i* on the basis of bad evidence; hence, *S* does not know that *i*. However, causal origination is not necessary for belief-basing. For *S* may cease to believe that *c* yet still retain the belief that *i* (*Sbi*) because he still believes that *e*. In this case, *Sbe* causally sustains *Sbi*. As a result, *S continues to believe* that *i* on the basis of conclusive evidence; hence, *S* knows that *i*. A case of causal overdetermination occurs if each of *Sbc* and *Sbe*, by itself, causally originates or sustains *Sbi*. Though this case is more controversial, some causal theorists claim that *S* knows that *i* since *S* believes that *i* partially on the basis of conclusive evidence.

The criterion for causal origination, sustainment and overdetermination is typically what explains, at least partially, the presence of a belief. Once this is made explicit, counterexamples abound for providing the motivation to go beyond these types of causal relations. For example, if *Sbc* causally sustains *Sbi* so that *only Sbc* explains the presence of *Sbi*, then *Sbe* bears no causal relation to *Sbi*. However, one may claim that *Sbe* would causally sustain *Sbi*, if *Sbc* were not to do so. If this counterfactual is true, then there is a conditional (or non-actual) causal relation between *Sbe* and *Sbi*. Marshall Swain calls this pseudo-overdetermination and for this reason claims that *S* believes that *i* on the basis of conclusive evidence; hence, *S* knows that *i*.

It is at this point that Lehrer's gypsy lawyer counterexample should be considered. For it is designed to be a case of knowledge which is devoid of any appropriate conditional causal belief-basing relation.[1] The gypsy lawyer believes that the testimony of his tarot cards that his client is innocent is true (*Sbc*). This causes him to believe that his client is innocent (*Sbi*), to review the evidence, and to compose a complex, though valid, line of reasoning which causes him to believe that there is conclusive evidence for his client's innocence (*Sbe*). However, powerful emotional factors sustain others' beliefs that his client is guilty. Lehrer concludes his case as follows:

> The evidence is quite conclusive ... but even he would find himself unable to believe his client could be innocent ... were it not for the fact that the cards told him it was so ... were his faith in the cards to collapse, then emotional factors which influence others would sway him too. Therefore the evidence which completely justifies his belief does not explain why he believes as he does, his faith in the cards explains that, and the evidence in no way supports, reinforces, or conditionally or partially explains why he believes as he does. (*Knowledge*, O.U.P., 1974, p. 125).

Lehrer's crucial claim is the denial of the following counterfactual: *Sbe* would causally sustain *Sbi*, if *Sbc* were not to do so. As a result, *Sbe* bears no conditional causal relation to *Sbi*. If this is a genuine case of knowledge, then Swain's account is incorrect. Interestingly, Harman concedes this point to Lehrer; see *Thought*, Princeton U.P., 1973, pp. 31-32.

However, Swain believes that his account can handle the gypsy lawyer's knowledge. For he claims the following:

> ... it is plausible to say that S would be caused to believe that his client is innocent partly by his belief that the evidence is conclusive if he were to have this latter belief and did not have the belief about the cards. (op. cit., 1981, p. 91)

---

[1] Contrary to what Kvanvig claims (op. cit., p. 153), the gypsy lawyer case does not motivate the move to a non-actual causal account. Rather, it is a counterexample to this move. See Lehrer, op. cit., 1971, p. 311, id. *Knowledge*, p. 124, and Swain's acknowledgement in his earlier 'Justification and The Basis of Belief' in *Justification and Knowledge*, G. Pappas ed., D. Reidel 1979, pp. 25-49, see p. 48, fn 21.

Curiously, Swain affirms the same counterfactual which Lehrer denies. As a result, he undermines the critical point of the gypsy lawyer case. However, he offers the following aside:

> On the other hand, if the belief that the evidence is conclusive is not at least a pseudo-overdeterminant of the belief in innocence, I see no ground for claiming that the gypsy lawyer has knowledge. (ibid.)

We simply have *two* distinct gypsy lawyer cases. Swain claims that the gypsy has knowledge only in his case. He has cleverly altered Lehrer's case from a counterexample to a support for conditional causal belief-basing.

Kvanvig claims that causal theorists have had two sorts of reactions to the gypsy lawyer case: either to deny the knowledge claim or to offer an account of conditional causal belief-basing (op. cit., p. 154). We can now see that this claim hides an ambiguity. For these reactions result from considering two distinct gypsy lawyer cases. To Lehrer's original case, causal theorists have had only one reaction: to deny that it is a case of knowledge.

II

Causal theorists may have the advantage vis-à-vis Lehrer's gypsy lawyer case. On the one hand, causal accounts of belief-basing fit into an overall causal view of many philosophical concerns, e.g. perception, intentional action, memory, and reference. On the other hand, Lehrer's case is an intuitively difficult case of knowledge. An argument like the following might be forthcoming from a causal theorist. If the gypsy lawyer knows that *i*, then his knowledge is based on the conclusive evidence. Consider the following necessary condition for *S*'s knowing that *i* on the basis of believing that *e*: *Sbe* makes it 'objectively likely' that *Sbi is true*. In other words, evidence makes it likely that one's beliefs are true. This is satisfied; for the gypsy's evidence is conclusive. Now let's alter the condition as follows: *Sbe* makes it 'objectively likely' that *Sbi occurs*. In other words, evidence makes it likely that one believes truths. This condition is not satisfied. For although the gypsy properly appeals to the evidence, he believes that his client is innocent *only* because of the tarot cards. As a result, *Sbc* is necessary and sufficient for the occurrence of *Sbi*. Consequently, *Sbe* has no effect on the likelihood of the occurrence of *Sbi*. It is 'fortuitous' that the gypsy believes what is true; so, he lacks knowledge. Swain's gypsy lawyer case is different in just this lacks knowledge. Swain's gypsy lawyer case is different in just this respect. *Sbe* conditionally causally sustains *Sbi*. Consequently, *Sbe* is sufficient for the occurrence of *Sbi*: *Sbe* makes it 'objectively likely' that *Sbi* occurs. Both of the above conditions are fulfilled in Swain's case.

If we accept the above argument, or some equivalent, then the causal theorist must offer some account of conditional causal belief-basing. It is at this point that we should take Kvanvig's criticisms seriously. He offers a series of interesting examples relevantly similar to Swain's gypsy lawyer case in order to show that the conditional causal belief-basing relation is vacuous.

Let's examine Swain's account of pseudo-overdetermination.

> Where *Sbc* causes *Sbi*: *Sbe* causally pseudo-overdetermines *Sbi* iff *Sbe* does not cause *Sbi* and if *Sbc* were not to cause *Sbi* and *Sbe* and *Sbi* were still to occur, then *Sbe* would cause *Sbi*. (Compare Swain 1981, p. 70)

This is consistent with Kvanvig's S1 (op. cit., p. 154). Kvanvig then notes that pseudo-overdetermination may be *n*-adic, whereas Swain's explicit definitions are formulated only in terms of dyadic overdetermination. The problem is then how to generalize the above account. Contrary to Kvanvig's claim (op. cit., pp. 154-5, fn 1), Swain is not unaware of the need for this move. He writes as follows:

> This discussion focuses on cases of dyadic overdetermination; however, the results will be stated in definitions that are fully general with respect to *n*-adic cases. (op. cit. 1981, p. 65)

Unfortunately he does not make it clear how this is so. Kvanvig makes such an attempt in his S3 (op. cit. pp. 155-6). The crucial notion is that of a set of counterfactual causes which allows for *n* members. Kvanvig characterizes set membership in the following manner. To claim that *Sbe* is a counterfactual cause of *Sbi* is to claim

> ... that were the belief [*Sbi*] and the appreciation of the evidence [*Sbe*] to occur and were nothing else to have caused the belief [*Sbi*], then the appreciation of the evidence would have caused the belief. (op. cit., p. 157).

(this is very close to the language Swain uses to handle the gypsy lawyer case, especially the use of the phrase 'nothing else', op. cit., 1981, p. 91).

Kvanvig's notion of a counterfactual cause is intended to be identical to Swain's notion of a causal pseudo-overdeterminant. Following Kvanvig's usage, we can generalize this notion as follows:

> Where *Sbc* causes *Sbi*: (*p*) (*Sbp* is a counterfactual cause of *Sbi* iff *Sbp* does not cause *Sbi* and if *Sbp* and *Sbi* were still to occur and nothing else were to cause *Sbi*, then *Sbp* would cause *Sbi*).

This condition allows for *n*-adic overdetermination. The key phrase 'if nothing else were to cause *Sbi*' rules out the actual cause and any counterfactual cause other than *Sbp* which might cause *Sbi*. The idea is to isolate *Sbp* and *Sbi* where nothing else but the former could cause the latter in order to determine if it would do so.

According to Kvanvig, this condition succeeds all too well. He claims that it

> ... could escape being a necessary truth only if it is possible that the belief in question [*Sbi*] be uncaused, for if the belief cannot be uncaused and if nothing else but the appreciation of the evidence [*Sbe*] is a cause, then the appreciation of the evidence *must* be a cause. (op. cit. p. 157)

As a result, the notion of counterfactual cause is vacuous. For take any logically possible cause (which is not also an actual cause) of *Sbi*, if it and *Sbi* were to occur and nothing else were to cause *Sbi*, then it would do so; hence, it is a counterfactual cause. Initially, it may be suspected that Kvanvig has simply confused the notions of counterfactual cause and logically possible cause. For the former implies the latter but not vice versa. The above condition seems to be a necessary truth only if we take 'nothing else' as 'no other logically possible cause'. Instead, if we take it as a more restrictive 'no other counterfactual cause', then its truth seems to be a contingent matter. But Kvanvig's insight seems to be that the notion of counterfactual cause, if formulated generally enough for epistemological purposes, reduces to the notion of logically possible cause. Kvanvig's challenge is to offer a unique account for the notion of counterfactual causation.

I believe that there are two ways to accomplish this task. The first involves a further specification of the set of counterfactual causes as an ordered series. Set membership may be defined partially as follows:

> where *Sbc* causes *Sbi*: (*p*) (if *Sbp* is a member of set *C* of counterfactual causes of *Sbi*, then *Sbp* does not cause *Sbi* and if *Sbc* and all members prior to *Sbp* (if any) were not to cause *Sbi* but *Sbp* and *Sbi* were still to occur, then *Sbp* would cause *Sbi*).

Kvanvig might reply that since any logically possible cause, other than *Sbp*, is such that if it were to cause *Sbi*, then *Sbp* would not do so, then all other logically possible causes must be included as members of *C* prior to *Sbp*. However, this reply contains an error. To see this consider the notion of causal overdetermination. Where some *Sbc* and *Sbl* causally overdetermine *Sbi*, then both *Sbc* and *Sbl* cause *Sbi*: we shall call them joint causal overdeterminants of *Sbi*. There are two properties of logically possible causes which distinguish them from counterfactual causes. One, any merely logically possible cause, e.g. *Sbl*, can be a joint causal overdeterminant with the actual cause, *Sbc*, such that both cause *Sbi*. Two, *Sbl* and any counterfactual cause, e.g. *Sbq*, would be joint causal overdeterminants of *Sbi*, if both were to cause *Sbi*. However, counterfactual causes cannot be a joint causal overdeterminant with the actual cause or any prior member(s). For by the above necessary condition, if *Sbp* is a counterfactual cause of *Sbi*, then it cannot also cause *Sbi*; it cannot be a joint causal overdeterminant (with *Sbc*) of *Sbi*. Furthermore, if *Sbp* has a prior member, *Sbq*, then if *Sbq* were to cause *Sbi*, then *Sbp* would remain a counterfactual cause and as such could not be a joint causal overdeterminant (with *Sbq*) of *Sbi*. This latter requirement can be made explicit as follows:

> (*p*) (if *Sbp* is a member of set *C* of counterfactual causes of *Sbi*, then if any member prior to *Sbp* were to cause *Sbi*, then *Sbp* would not do so).

I believe that the above two necessary conditions for counterfactual causation suitably restrict the notion so that not all logically possible causes are counterfactual causes. Consequently, the notion of counterfactual causation is not vacuous and may be used profitably in an account of belief-basing.

Alternatively, one could respond to Kvanvig by noting an oddity emphasized in the antecedent of the following:

> if *Sbc* were not to cause *Sbi* and both *Sbe* and *Sbi were still to occur*, then *Sbe* would cause *Sbi*.

In order to deny this counterfactual, one must construct a case which satisfies the antecedent but not the consequent. The result is a case where *Sbi* occurs and is uncaused or caused by something other than *Sbe*. However, Lehrer's gypsy lawyer case does neither. Lehrer claims that if the cards were no longer to cause the gypsy's belief in his client's innocence, then he would no longer have this latter belief – regardless of his belief in the evidence. We can express this as follows:

> if *Sbc* were not to cause *Sbi* and *Sbe* were still to occur, then *Sbi would not occur*.

Lehrer's case does not deny the earlier counterfactual, for it violates the supposition that the gypsy retains belief in his client's innocence. Indeed both Swain and Kvanvig state the conditions for causal belief-basing in this manner, thereby restricting possible counterexamples.

This allows us to formulate a simplified account of counterfactual causation. We are even able to return to the unrestricted use of 'nothing else' in the following:

> if *Sbc* were not to cause *Sbi* and *Sbe* were still to occur and nothing else were to cause *Sbi*, then *Sbe* would cause *Sbi*.

If this antecedent is satisfied, then it remains an open question whether *Sbe* would cause *Sbi*. Kvanvig's claim that it *must* do so depends crucially upon the stipulation in the antecedent that *Sbi* still occur. Once this is removed, the difficulty vanishes.

Keith Lehrer in 1971 proposed the gypsy lawyer case and advanced an alternative non-causal account of belief-basing. The latter has not been widely accepted. As a result, Lehrer's critical case rests on the difficult intuitions of knowledge for his gypsy lawyer. I submit that Kvanvig's case fares no better. For he claims that no account of conditional causal belief-basing is adequate for

Swain's gypsy lawyer case. I have concluded that his claim is incorrect. The critic of a causal theory is then left only with Lehrer's original case and this is not very satisfying. Consequently, the task for this critic is as yet unfulfilled.[1]

The University of Nebraska-Lincoln,
Lincoln, Nebraska 68588, U.S.A.

©LORY LEMKE 1986

---

[1] I wish to thank Al Casullo, Pierrana Garavaso, Ted Uehling, the Editor, and a referee of this journal for helpful comments on this paper.

ACTING INTENTIONALLY AND MINIMAL ABILITIES

By MICHAEL J. COSTA

COMPARE the following two cases:

Case I. Smith is offered ten dollars if he can shoot a bottle placed on a distant wall. Smith is not an expert marksman and he knows it. The chances that he will hit the bottle at this distance are remote (say, one in a hundred). Yet Smith has nothing to lose. He takes aim as carefully as he can and fires, hoping to hit the bottle. The shot hits the bottle, and Smith gets the ten dollars.

Case II. Smith is playing a board game. He needs to roll a six with the die to win the game, thereby winning ten dollars. He throws the die, hoping to throw a six. The die is a fair die and Smith has no special manipulative powers with dice, yet he does throw a six and wins the ten dollars.

The cases resemble those in a problem posed in ANALYSIS some years ago (Problem No. 16, 37.3, June 1977, p. 97) which produced substantial subsequent discussion (38.3, June 1978, pp. 113-18, 40.3, June 1980, pp. 113-18, 41.3, June 1981, pp. 156-9, 41.4, October 1981, pp. 216-18). My concern here, however, is different from that emphasized in the discussion of the earlier problem (which was, roughly, whether an intentional action can be generated by an act that is purely a matter of chance).

I think that intuitions are fairly clear that Smith shot the bottle intentionally in case I, but that Smith did not throw a six intentionally in case II. Case I shows that one can do something intentionally without having the ability (in any ordinary sense) to do it. Smith does not have the ability to hit targets with any consistency at that distance. But then why isn't Smith's action intentional in case II? Smith's actual chances of succeeding were much better in case II (one out of six) than they were in case I (one out of one hundred). He didn't have the ability to do either act. What is the relevant difference between the cases that explains why Smith's action is intentional in case I and not in case II?

The answer, I think, is that while having the ability to do an act is not a *sine qua non* of doing that act intentionally, having the *minimal ability* to do the act is. Roughly, one has the minimal ability to do an act if the *chances* of one's doing the act are *improved* by one's trying to do it. Smith's chances of throwing a six were not improved by his trying to throw a six, but Smith's chances of hitting the bottle were improved by his trying to hit the bottle with the shot.

This rough account of minimal ability is subject to serious objections, however. If one is a nervous sort of individual, it may be that one's chances of doing something *decrease* when one *tries* to do it. *Trying* is not quite the concept that is needed. There is a further problem. Suppose that the game in case II were such that Smith had some option as to whether to throw the die (he could have passed his turn). If Smith had not tried to throw a six, he wouldn't have thrown the die. Now Smith's chances of throwing a six are much better if he throws the die than if he doesn't (the latter chances being zero), so Smith's trying to throw a six does improve the chances of throwing a six. The difficulty, of course, is that in formulating the rough account we were taking as a given that the agent performed the 'means' act. Given that Smith performs an act of throwing the die, his doing it in order to throw a six has no impact whatsoever on his succeeding in throwing a six; whereas, given that he performs an act of firing the gun, Smith's chances of hitting the bottle are improved if he does it in order to hit the bottle.

These difficulties can be avoided by formulating the notion of minimal ability a bit more carefully:

S has the *minimal ability* to do an A if and only if:

(1) S has the ability to do an A, or
(2) S has the ability to do a B ($\neq$ A) such that $P(A/B(A)) > P(A/B-(A))$.

The symbolization requires some commentary. The expression '$P(A/B(A))$' is to be read: the probability that S does an A given that S does a B in order to do an A thereby. The expression '$P(A/B-(A))$' is to be read: the probability that S does an A given that S does a B but not in order to do an A thereby.

In relying on the notion of doing a B in order to do an A thereby the account avoids the implication of 'trying' that special effort or premeditation is involved. And by bringing in reference to the 'means' act, we avoid the other problem noted above. In case II Smith does not have the ability to throw a six, and there is no act he has the ability to do which is such that the probability that he throws a six given that he performs that act in order to throw a six thereby is greater than the probability that he throws a six given that he performs that act but not in order to throw a six thereby. In case I, on the other hand, while Smith does not have the ability to shoot the bottle at that distance, it is the case that there is some act that Smith has the ability to do, namely, aim and fire the gun, which is such that the probability that he shoots the bottle given that he performs an act of that type in order to shoot the bottle thereby is greater than the probability that he shoots the bottle given that he aims and fires the gun without doing it in order to hit the bottle thereby.

The value of using the notion of doing a B in order to do an A thereby is further revealed by considering another kind of objection. Suppose it is claimed that there *is* an act that Smith can do that will improve his chances of throwing a six, namely, Smith can drill and weight the die. Presumably his drilling and weighting the die in order to throw a six makes it more likely that he throws a six than if he drills and weights the die but not in order to throw a six. It may seem, therefore, that clause (2) is satisfied. But two responses are available. First, the act that Smith does not have the minimal ability to do is to throw a six with a *fair* die. Smith's drilling and weighting the die do not improve his chances of throwing a six with a fair die. Second, Smith would not drill and weight the die in order to throw a six *thereby*. Drilling and weighting the die is a *preparatory* act to throwing a six, not an act that generates the throwing of a six. One may walk to a door in order to open it (or should we say, more strictly, in order to put oneself in position to open it?), but one does not walk to a door in order to open it *thereby*. The act B in question must be one that generates the act A in order to satisfy clause (2).

The analysis explains minimal ability partially in terms of ability, but this should not produce explanatory circularity so long as explaining the notion of ability does not presuppose the notion of minimal ability. I shall assume that this is the case, without making any attempt to engage in the difficult task of analysing ability. *If* we can understand the notion of ability, my account provides the basis for understanding the notion of minimal ability in terms of it.

Clause (1) seems appropriate (we should have the minimal ability to do anything we have the ability to do), and it is needed in any event to handle the case of basic actions, which would not qualify under clause (2). Also it handles the case of actions that are the inevitable results of other acts that one has the ability to do. Consider the action of displacing air. This act does not qualify by clause (2). The probability that I displace air given that I move my hand in order to displace air thereby is no better than the probability that I displace air given that I move my hand but without moving it in order to displace air thereby. But, since I have the ability to displace air, I also have the minimal ability to do it by clause (1).

I do not claim that *mention* of minimal abilities is required in any adequate account of acting intentionally. It may be, for example, that acting intentionally can be explained in terms of one's action satisfying or matching what one wills to do. If one lacks the minimal ability to do an act, that is because one cannot *will* to do that act (though one can desire or hope to do it). The actual presence of a volition to do an A reveals itself in having some impact on the probability of the agent's succeeding in doing an A. Carrying through on this suggestion, however, is the subject for another project.[1]

University of South Carolina,
Columbia, S.C. 29208, U.S.A.

©MICHAEL J. COSTA 1986

---

[1] I have benefited from discussions on this subject with Davis Baird, Barry Loewer and Ferdy Schoeman.

PROSTHESIS AND PRE-EMPTION

By JAMES MONTMARQUET

CHRISTOPHER Peacocke has worried (*Holistic Explanation*, Oxford, 1979, p. 87) over the case of the 'knowledgeable neurophysiologist'

(1) who decides on a particular occasion to produce in me exactly the motor impulses to realize what he knows, from my neurophysiological states, to be my intentions.

His worry, at bottom, is that the physical realization of these intentions would (intuitively) not count as *his* intentional actions – even though they would satisfy Peacocke's criterion of intentional action (roughly: that one's X-ing behaviour be caused by an intention to do X, where both cause and effect are describable in such a way that the laws connecting them specify the one as a *function* of the other). The remedy for this proposed by Peacocke is specifically to exempt (ibid., p. 88) causal chains including the intentions of another person as possible intentional actions. On my own diagnosis of this case ('Causal Deviancy and Multiple Intentions', ANALYSIS 42.2, March 1982), Peacocke's concern is somewhat misplaced, for we can easily imagine cases in which such intervening intentions of another agent do *not* defeat attributions of intentional agency, for example:

(2) S's brain has been wired to T's in such a way that when he forms an intention to do X, this will cause T to form the

> same intention (that S do X), the energy from which will be supplied to close what would otherwise be a gap in the causal chain connecting S's intention and his X-ing behaviour.

What ultimately distinguishes such cases from (1), on the account I proposed, was that in the latter (but not the former) the occurrence of the second agent's intention *itself* marks the start of an intentional action.

In a more recent publication John Bishop, whose earlier discussion of Peacocke ('Peacocke on Intentional Action', ANALYSIS 41.1, March 1981) had much inspired my own, takes strong issue with this approach to cases typified by (1) and (2), and instead endorses an alternative approach considered and rejected by me in the above (see 'Causal Deviancy and Multiple Intentions: A Reply to James Montmarquet', ANALYSIS 45.2, June 1985). Bishop's main charge is that once we grant, as of course I have, that an intentional action can be accomplished through the 'prosthetic mechanism' of another's intentions, then it is 'purely arbitrary' (p. 165) to disallow as intentional actions cases in which another's intentions mark the start of an intentional action of his own. On the alternative view of (1) proposed by Bishop, the decisive factor is simply that the neurophysiologist 'pre-empts' the agent from even attempting to realize his intention:

> For an agent performs a basic intentional action only if his forming (or having) the relevant basic intention still *leaves it up to him* to carry out that intention or not. And, where pre-emption occurs, it is no longer up to the agent whether his basic intention is matched (p. 167).

Now I have, as Bishop acknowledges (p. 168), admitted that in (1) the original agent has been 'pre-empted' from acting; the issue here is somewhat subtle. For me, an agent can be pre-empted, if one wants to use that term, from X-ing even *after* he has begun to act on his intention to X and even if he completes the relevant bodily movement for X-ing. He will be pre-empted whenever he completes his behaviour only as a consequence of someone else's intentional act of making him do X. For Bishop, such an intervention – if begun after 'recognizing the agent's intention *and the cortical event initiating his action*' (p. 167, Bishop's emphasis) – would not rule out the first agent's having intentionally X-ed. Thus, Bishop would allow that the patient does intentionally raise his arm in such a case as this:

> (3) A patient who is recovering from a stroke is trying to regain full control of his limb movements. In another room a neurologist is monitoring the patient's progress. When he notices, on one occasion, that the patient is not going to be able to carry out his basic intention to raise his arm, the doctor intervenes (employing an elaborate set-up) so as to make the patient's arm go up, this intervention having begun

> *subsequent* to the point at which the patient has begun to attempt to carry out his intention.

The first point I want to make about this case is that it is surely implausible to characterise it as one in which the patient has successfully performed a *basic* intentional act of raising his arm. To be sure, we may admit that the patient has done *something* – perhaps 'has tried to raise his arm' or 'has done something which causes someone else to raise it' – but it is the *neurologist*, not he, who has raised the patient's arm. (For the sake of the patient's confidence, the neurologist may well wish him to *think* that he has raised it; indeed the neurologist no doubt *wishes* that the patient himself had been able to do this – but in plain truth he has not.) Or consider this: if we suppose that both the patient and the neurologist have performed basic intentional acts of raising the patient's arm, then, since the acts have clearly begun at different times (the patient's beginning slightly before the doctor's), this case would be correctly described as one in which 'first the patient began to raise his arm, then the doctor began to raise it' – which conveys very much the wrong idea of what has happened. (Compare the very much more natural: 'first the patient tried to raise his arm, then the doctor raised it (for him)', or 'first the patient did something which got the doctor to raise his arm, then the doctor raised it'.) There have not been *two* successful raising of this arm – one done by one agent and one by the other – nor one successful arm raising with joint agency (though this would have been the case had the patient applied *effectual* effort during the doctor's intervention). Rather, there has been *one* successful arm raising whose agent has been the doctor (and one unsuccessful attempt to do this as a basic act – on the issue of non-basic arm raisings, see below).

The relation of the parties in (2) and (3), then, is not the same. In (2) the second person serves merely as a bearer of ('transeunt') causality by which the first agent's powers are fortuitously extended. In (3), however, the second and intervening person – by *taking* control of the situation – makes himself anything but one who extends the causal powers of another. (Granted, there are cases in which we would correctly describe the neurologist as extending the patient's power to control events; but these would be ones in which the patient literally controls the behaviour of the neurologist – say, through prior hypnosis. To the extent, though, that such control is exerted on the doctor, his behaviour becomes that much more dubiously classified as *his* intentional action. Thus we come back to the importance of whether a chain includes the intentional act of another.)

Our discussion thus far has focused on the analysis of intentional *basic* action, but it may be thought that Bishop can fall back to defend the following thesis concerning *non*-basic action: that at least in those cases in which it was S's plan to do X by means of another agent's more direct agency, S's doing X is an intentional non-basic act. Indeed one of Bishop's complaints against my earlier discussion is that:

> Succeeding in carrying out an intention that someone else do something for one is not less than a 'full-fledged' intentional action, as Montmarquet implies (p. 109): it is simply a species of *non-basic* action – one which can be achieved only by first executing other more immediately achievable intentions (pp. 164–5).

Well, suppose we alter case (3) to make it part of the original agent's plan that the neurologist intervene in the way that he does. What effect ought this to have on our description of the case?

Let us begin by contrasting three descriptions of what the patient in (3) has done:

(i) brought about certain events in his brain.
(ii) gotten the neurologist to raise his (the patient's) arm.
(iii) raised his arm.

Now (i), let us assume, constitutes a basic description of the patient's action (if you prefer, substitute 'tried to raise his arm' or some other favoured description); but (ii) is obviously a non-basic description – which brings out an important point apparently in support of Bishop's view, namely: an act can be redescribed not only in terms of such ordinary circumstantial factors as determine that a flicking of a switch will be a turning on of a light, but also in terms of factors under the voluntary control of another agent – and even where, as in (3), at the time of the original action it is still up to that other agent whether he will intervene. Thus, as Bishop would want to insist, there can be non-basic acts whose completion depends on the exercise of someone else's agency. What objection, then, can there be to a similar construction of (iii): that it is a redescription of the patient's basic act (say, of making certain events happen in his brain) in the light of the fact that this basic act caused someone to raise the patient's arm? I would answer by inviting a comparison of these:

(a) By initiating certain brain events, the patient got the neurologist to raise his arm.
(b) By getting the neurologist to raise his arm, the patient raised his arm.
(c) By initiating certain brain events, the patient raised his arm.

Now it seems to me that whereas (a) is a perfectly acceptable description of what has happened, (c) is at best a grossly misleading one and (b) is *odd* – in just the way that the claim:

By getting someone to shoot Kennedy, Castro shot Kennedy.

is so. But if we can, as in (a), redescribe an act in terms of the subsequent efforts of another agent, why cavil at (b) and (c)? The answer is that there is no objection to such redescription when *and only when* they accurately and nonmisleadingly describe the situation by making clear *who the controlling (or proximate) agent is*. Obviously (a) does this by making it clear that the controlling agent with respect to the event of the arm's rising is the neurologist. By contrast, (c) is misleading because it conveys the idea that the patient, not the doctor, exerted the controlling influence on this event. Finally, (b) is odd – rather than misleading – because it conveys the idea that the patient both was and was not the controller. The proposal, then, to treat the raising of an arm (under the circumstances in question) as a non-basic action has a serious conceptual difficulty surrounding it – which, since it seems to be an entirely conceptual proposal to begin with, puts it in rather dire straits.

Let me turn now to assessing the significance of all of this for the larger project – of primary interest to Peacocke, Bishop and myself – of giving an account of intentional action. In his earlier submission, Bishop had argued that Peacocke's blanket exemption of all causal chains containing another agent's intention as intentional acts was *ad hoc*. To this I replied that since not all such chains had to be exempted but only ones in which the intervening intention itself marked the start of an intentional act, Peacocke had a way round Bishop's objection. For if his account of intentional action is *otherwise* successful we could apply it recursively to the problem of second agent interventions, roughly to this effect:

> S's doing X is an intentional basic action of S's if and only if S's X-ing is 'functionally explained' (in Peacocke's sense) by S's intention to do X, and the causal chain connecting the latter to S's X-ing does not contain as a proper part any other intention to do X which itself functionally explains S's X-ing.

Here the original restriction appears as no longer *ad hoc*, but simply as a cyclical application of the functional explanation requirement itself. Now, given all of this, Bishop rightly takes his later paper to cast doubts on the acceptability of the above criterion and to once again open up the prospect (cf. p. 168) that an acceptable account of intentional action perhaps may not be formulable entirely in terms of event-causation. If my own arguments here have been plausible, however, Bishop is doubly in error: first (and here, ironically so, given his own doubts about an event-causation account of intentional action) because he has failed to notice that not taking seriously the difference between cases (2) and (3) amounts to a failure to acknowledge a distinctive difference between human agency and ordinary transeunt causality; second, because if I have been right that this distinction needs to be preserved (unaltered), there remains no important objection to an event-causation analysis of intentional action – or none emanating from the problem of intervening intentions!

Washington State University,
Pullman, Washington 99164-5130,
U.S.A.

©JAMES MONTMARQUET 1986

WAYS OF SPEAKING OF IMAGINATION

By ALAN R. WHITE

IN 'Speaking of Imagination' (*Language, Mind and Brain*, edd. T. W. Simon and R. J. Scholes (1982), pp. 35-43, reprinted with additions as Chapter III of *The Matter of Minds*, Oxford 1984), Zeno Vendler advances several linguistic arguments for the thesis that all that can be imagined are experiences. I shall try to show that these arguments are invalid. My aim in this is to get nearer to a correct understanding of the nature of imagination by throwing doubt on this restriction. What Vendler has done is to generalise about imagination from a selection of examples like 'imagine swimming' and 'imagine someone swimming'.

Vendler begins by distinguishing between 'Imagine Ving', for example 'imagine swimming, eating a lemon, being on the rack, whistling in the dark, seeing or hearing so and so', and 'Imagine A (or A's) Ving', for example 'Imagine yourself swimming, running or whistling, him swimming, running or whistling, Rubinstein playing'. He then argues that the second group can be reduced to the first on the ground that 'imagine A (or A's) Ving' is elliptical for 'imagine seeing, (or hearing) etc. A (or A's) Ving'. Thus, to imagine yourself or him (his) swimming or playing is, he holds, to imagine seeing yourself or him (his) swimming or hearing yourself or him (his) playing. Hence, all such imagination calls for 'the representation of the experiences', whether tactual, muscular or kinaesthetic sensations in the first group or visual and auditory in the second group, which the imaginer would have if he were in the imagined situation.

A first objection to this is that Vendler has been misled by his restricted diet of examples. To imagine sacrificing everything for one's principles, going to all that trouble just for the sake of a dead cat, selling one's birthright for a mess of pottage, working all one's life for a pittance, paying £1,000 for a hat, etc. is not to give oneself a representation of any experiences. Nor is imagining oneself or someone else doing any of these or imagining, for example, him (or his) running a university, working with a woman, agreeing to this, changing his mind about that, approving of so and so or submitting to such and such linked to any representation of experiences either directly of a tactual, muscular or kinaesthetic kind or indirectly of a perceptual kind. Many of the situations in which I can imagine myself being, for example worried by this, inclined to do that, undecided between these two alternatives, misunderstood by my friends, are not situations in which I imagine seeing myself.

A second objection is to Vendler's thesis that the second group of imaginings, that is 'imagining A (or A's) Ving', is reducible to the first, that is, 'imagining Ving', on the ground that 'imagining A (or A's) Ving' is elliptical for 'imagine seeing (hearing, etc.) A (or A's) Ving' (by contrast Ryle, in *The Concept of Mind*, p. 273, says that what I imagine is always myself Ving). One reason for this false thesis is again a restricted diet of examples. To imagine or to be unable to imagine either oneself or another('s) agreeing to this, changing one's mind about that, approving of so and so or disapproving of such and such, sacrificing all for love, spending so much money on a hat, etc. is not to imagine or be unable to imagine seeing (or hearing) oneself or another('s) doing this.

Even in sensory examples – to take a few modelled on his own – there is a difference, neglected by Vendler, for example, between imagining seeing a lion coming towards one and imagining a lion mating with a tigress, between imagining seeing a building towering above one and imagining a building being venerated by a tribe, between imagining hearing church bells far away or the sound of a train coming close and imagining church bells splitting in twain or the sound of a train's annoying the passengers, and between 'It is myself (and not you) that I imagined seeing swimming in that water' and 'It is myself (and not you) that I imagined swimming in that water'. It would be quite mistaken to take the second member of each of these pairs, which are examples of imagining A (or A's) Ving, as examples of imagining seeing or hearing A (or A's) Ving. Neither to imagine being shaved by a monkey nor to imagine yourself or someone else being shaved by a monkey is to imagine seeing yourself or someone else being shaved by a monkey. Equally, 'Fancy you (your) being here' is quite different from 'Fancy seeing you here', which Vendler, who believes that 'fancy' is a 'near synonym to imagine', ought to think its equivalent. 'Fancy Caruso singing that aria' is not the same as Vendler's example 'Fancy hearing Caruso singing that aria'.

Vendler's thesis that to 'imagine A (or A's) Ving' is to 'imagine seeing (hearing, etc.) A (or A's) Ving' forces him to disallow B's imagining A (or A's) seeing or hearing something on the plausible grounds that one cannot see or hear another('s) seeing or hearing something, as contrasted with him (his) looking at or listening to it. But there is nothing strange about my imagining Shelley('s) hearing a skylark and being disappointed by it or my imagining Othello entering his bedroom and seeing Desdemona with Cassio.

The difference, I submit, between 'imagining Ving' and 'imagining A (or A's) Ving' is, as the very grammar suggests, simply that between emphasising the Ving itself as the object of imagination and emphasising the agent of Ving as the object. Indeed, Vendler himself really takes this view too ('in complying with (1) [sc. "imagine swimming"] you imagine doing something, in complying with (2) [sc. "imagine yourself swimming"] you imagine yourself (or, in other cases, somebody else) doing something'). The more subtle difference between 'imagining A Ving' and 'imagining A's Ving' – for example 'imagine (can you imagine, I can't imagine) me (my) doing a thing like that'; that is the difference between the participial and the gerundive Ving (a distinction doubted by F. R. Palmer, *A Linguistic Study of the English Verb*, 1965, ch. 9) – is a difference in emphasis on the agent and on the agent's engagement. Contrast 'Imagine him running a University' and 'Imagine his overlooking a simple point like that'.

In other sensory examples, Vendler seems to me to have misinterpreted the element of perception. He argues that, for example, 'Imagine the battlefield from above, the statue sideways or the music coming from a distance' are elliptical for 'Imagine seeing the battlefield from above and the statue sideways or hearing the music from a distance' on the ground that one cannot, for example, imagine a thunderbolt from the side or the taste of lemons from above or the Escorial from left to right. Certainly, qualifications such as 'from above, sideways, from a distance' etc. can only go with what is seen or heard; but this is not sufficient to show that to imagine these is to imagine seeing or hearing these. To imagine, for example, the battlefield from above, the statue sideways or the music from a distance could be, and I think is, to imagine these as seen or heard in this way. But to imagine X as seen or heard in a certain way is not to imagine seeing or hearing them in this way, but to imagine what they would look or sound like if seen or heard in this way.

Similarly, Vendler's thesis that to imagine X, for example an elephant or a thunderbolt, is to imagine seeing or hearing X and, hence, that imagination here is 'restricted to the representation of objects with visual, auditory or otherwise sensory aspects' is also due to his restricted diet. One can imagine with ease or difficulty, for example, someone's anger, pleasure, surprise, reaction, etc. at the news, the trouble I went to in order to get that present for you, the dangers or difficulties of a certain course of action. Just as Vendler neglects the difference, noted above, between, for example, imagining seeing a lion coming towards me and imagining a lion mating with a tigress, so he neglects the difference between, for example, imagining the look of surprise on a man's face and imagining his surprise and, therefore, of imagining seeing an elephant and imagining an elephant. Hence, he also supports Berkeley's famous denial that one can imagine 'a tree or a house existing by itself, independent of, and unperceived by any mind whatsoever' because he, like Berkeley, equates 'imagining an unseen tree' with 'imagining seeing an unseen tree', though Vendler allows 'imagining an unobserved tree'. But though the description one gives of what one imagines when one imagines something visual may usually be in the same terms as a description of what one sees, it does not follow that to imagine something visual is to imagine seeing, or oneself seeing, something visual.

Vendler is further mistaken in asserting that 'one cannot imagine numbers, classes, human rights or moral virtues' because they are not perceptual. I can be asked to imagine a number without any fives in it, a class which contained no reference to itself, a human right with no corresponding duty, a moral virtue which is now despised. Incidentally, another way in which Vendler's over-emphasis on experience misleads him is in his declaration that one cannot imagine doing or being what one currently is. One can easily imagine Ving when one is Ving provided one does not think (or know) one is Ving, as when one imagines being abandoned in a cave when one has, unknown to one, been abandoned, or imagines being stalked by a wild animal when, unknown to one, one is being stalked.

Vendler also links what he calls 'objective' imagining, which he holds occurs both in 'imagining A (or A's) Ving' and in 'imagining X', with picturing. He says that 'imagine' in 'Imagine yourself swimming in that water' can 'be replaced by the verb *picture* without loss of meaning or grammaticality'. But to this one can object, first, that one cannot sensibly substitute 'picture' for 'imagine' in such examples of 'imagine A (A's) Ving' as 'Imagine Jones('s) spending all that money on a hat', 'Can you imagine me (my) marrying someone like that', 'I can't imagine you (your) agreeing to such a proposal'. Does Vendler want to say that there are two senses of 'imagine', a pictorial and a non-pictorial one? He nowhere suggests this. Secondly, Vendler's alleged synonymity of 'picture' and 'imagine' is inconsistent with the conclusion of two other theses of his. For if, as he holds, 'imagine yourself swimming' is 'nothing but an elliptical product' of 'imagine seeing yourself swimming', then the synonymity of 'picture' and 'imagine' should imply that 'picture yourself swimming' is elliptical for 'picture seeing yourself swimming'. But equally, if, as he also holds, 'picture' 'simply amounts to *imagine* seeing' (Ryle, op. cit., p. 256, held that 'picturing X' is 'imagining that one sees X'), then the synonymity of 'imagine' and 'picture' should imply that 'picture seeing yourself swimming' ought to amount to 'imagine seeing seeing yourself swimming', which is nonsense.

Vendler himself rejects the legitimacy of 'picture seeing A Ving'; but his reason for this is suspect. He says it is because *'picture'* simply amounts to *'imagine seeing'*. The implausibility of this can be brought out by the analogous example of 'see'. If, perhaps, 'I can't imagine him (his) giving someone £5 to hold his horse' has an analogy in 'I can't see him (his) giving someone £5 to hold his horse', then the former's alleged equivalent, 'I can't imagine seeing him (his) giving someone £5 to hold his horse' would have an analogy in 'I can't see seeing him (his) giving someone £5 to hold his horse'. It would be wildly implausible to explain this latter nonsense away by assuming that this is a use of 'see' which 'simply amounts to' 'imagine seeing'.

Vendler admits that 'imagine' is 'often used in a way which does not fit into the pattern sketched thus far'. But this admission is only made for 'imagine that' and 'imagine what' and it is coupled with two theses, namely that this other use of 'imagine' is linked to the non-sensory and that it is 'semantically quite similar to *suppose, think of, consider the possibility*, etc.'. To this I would object, first, that I have argued above that even 'imagine Ving', 'imagine A (or A's) Ving', and 'imagine X' do not fit Vendler's pattern. Secondly, that no evidence has been given – and I would argue at some other time that there is no evidence – for more than one sense of 'imagine'. Thirdly, 'imagine that' and 'imagine what' are as naturally and commonly used with the sensory as with the non-sensory, as when one is asked, for example, 'to imagine that you are on a palm fringed, coral beached, desert island with a gramophone and ten records' or 'to imagine what would happen if the sound of a fly's wings were magnified a thousand times'. Fourthly, that 'imagine' is never synonymous with 'suppose'. The assertion that it is is an extremely common fallacy perpetrated by a surprising number of postwar writers on imagination, against which I argue elsewhere. Here it is sufficient simply to raise two objections to Vendler's own examples. One is that though we can say both 'imagine' and 'suppose' that there is life outside the Solar system, we can also say, for example, 'I just can't imagine that there is life outside the Solar system', but not 'I just can't suppose that there is life outside the Solar system'. The other is that though we can say 'Imagine what would have happened if Goldwater had been elected President', one cannot say 'Suppose what would have happened if ... ', and though one can say 'I can pretty well imagine why he did it', one cannot say 'I can pretty well suppose why he did it'. (Curiously, Annis Flew, 'Images, Supposing and Imagination', *Philosophy* 28 (1953), pp. 246-54, who also suggested that there is one sense of 'imagine' equivalent to 'suppose', also uses the example, 'Imagine what would happen if ... ' without apparently realising that one cannot even grammatically here substitute 'suppose'.) These linguistic differences are partly due to the fact that imagining, but not supposing, is something one can try, succeed or fail at; something which can be easy or difficult, something one can be good or bad at. There are feats and flights of imagination, but not of supposition. People differ in their powers of imagination, but there are no powers of supposition.

The University, Hull © ALAN R. WHITE 1986

METAPHYSICAL REALISM AND HISTORICAL INTERPRETATION

By PETER SMITH

TO begin with a truism: opportunities for verifying singular historical propositions are normally evanescent. To take a trite example, suppose Jack is alone and rolling dice blindfold. He throws the dice once, fumbles on the table to retrieve them and throws them again. There is now no way in which he or anyone else can verify whether he actually rolled double six first time. There was a period between the two throws when, by peeking, Jack could have discovered how the dice fell. But since he kept the blindfold on, the possibility of finding that out is now lost for ever.

Here's another truism: interpretations of someone else's actions, even when backed by excellent evidence, can be erroneous. To take another trite example, suppose Jill praises Jane but we wrongly take her remarks as intended sarcastically. This interpretation best fits what else we know of Jill and her relationship with Jane. And when Jill protests that her remarks were indeed intended innocently, we say 'The lady doth protest too much, methinks!' and (again wrongly) we take her reaction as further proof of our interpretation.

Neither truism supports anything properly called 'metaphysical realism'. For example, someone who identifies truth with verifiability-at-some-time-or-another will just take the first truism as reminding us that the ideal moment for verifying propositions about the past is itself normally in the past. And even a more radical anti-realist who insists that truth cannot outstrip verifiability-now-or-in-the-future can still accept the proposition that a vast number of past events are beyond our epistemic reach, for *that* proposition is now warrantedly assertible if anything is. Which isn't to say that there are no problems for the anti-realist about the past, but only that the transience of evidence about the past is not in itself a damning difficulty. Likewise, the anti-realist (instrumentalist) about interpretation can equally happily accept that interpretations which are well grounded by everyday standards may yet fall short of the truth. For that is merely to acknowledge the defeasibility of interpretation, together with the fact that what counts as good (or even ideal) evidence by everyday standards is actually only part of the potentially available evidence – after all, we don't normally submit Jill to a lie-detecting test, or torture, or an intensive psychological test in order to elicit the truth. Which isn't to say that there are no problems for the anti-realist about interpretation, but only that the possibility of interpretations going wrong – even though well grounded on the available evidence – is not in itself an objection.

Further, the combination of our two truisms is, as far as I can see, no more potent than either of them taken separately. However, Norman Melchert dissents ('Metaphysical Realism and History', ANALYSIS 46.1, January 1986). He argues that our two truisms taken together establish nothing less than full-blooded metaphysical realism as understood by Hilary Putnam or J. J. C. Smart. Historical enquirers construct their interpretations on the best available evidence. But we will not always 'be able to recover or reconstruct certain facts relevant to the interpretation of... historical events' (opportunities for verification are fugitive); and the 'rational and justifiable' constructions we put on events on the basis of all the evidence we actually have or may come to have 'may still be false' (interpretation is fallible). Hence, says Melchert,

> The best evidence available, or *ever to be available*, may yet confirm a falsehood. And metaphysical realism is correct.

But the inference here is surely a mistake.

Metaphysical realism as understood by Putnam and Smart turns on the doctrine that *an ideally well supported general theory of the world may still be quite radically wrong in its overall structure and ontological presumptions.* Our experience systematically represents us as embodied inhabitants of a four-dimensional space-time world of material particulars. The assumption that the world is broadly speaking as it is thus represented as being is ideally well supported. But for all that, our experience could be delusory, says the realist attacked by Putnam; for we could be brains in a vat. Or, less wildly, perhaps the world is really *five*-dimensional, says the realist defended by Smart (see his 'Metaphysical Realism', ANALYSIS 42.1, January 1982, and cf. my reply in ANALYSIS 43.2, March 1983).

Now plainly there is no route from the premise (a) that the best available historical interpretation of some particular past events may be false to the conclusion (b) that we may be going radically wrong about the basic general character of the world. On the contrary, as we noted before, the premise combines two truisms which are as available to a thorough-going anti-realist as anyone; and such an anti-realist can consistently accept (a) while characteristically denying (b). Or to put it another way: the empirical realism expressed by (a) is quite consistent with a transcendental idealism which rejects (b). Hence Melchert's argument cannot settle the issue of metaphysical realism as standardly understood.

The University College of Wales,
Hugh Owen Building, Aberystwyth,
Dyfed, SY23 3DY

©PETER SMITH 1986

NOONAN ON NAMING AND PREDICATING

By E. J. Lowe

IN the course of arguing for a distinction between restricted and unrestricted quantification, Harold Noonan entertains the thesis that 'something can be predicatively a cat without being capable of being named by a name *for* a cat' (see his 'Relative identity: A reconsideration', Analysis 46.1, January 1986, p. 8). He explains the supposed distinction between names *for* F's and names *of* F's by saying that 'names *for* F's [are] names associated with the criterion of identity *same F*, and names of F's [are] names naming things which are predicatively F' (ibid., p.7). Further, Noonan speaks of something being 'predicatively' an F by supposed contrast with its being only 'constitutively' an F, in the sense of David Wiggins's 'is' of constitution. However, I fail to see how something which is *in*capable of being named by a name 'for' a cat can *be a cat* in anything like the same sense that something which *is* capable of being so named can, for reasons adumbrated in my article 'On being a cat' (Analysis 42.3, June 1982). As far as I can see, something is *capable* of being named by a name 'for' a cat just in case it is the sort of thing whose identity- and persistence-conditions are governed by the criterion of identity associated with the sortal term 'cat'. And to say that such a thing *is a cat* ('predicatively', if you will) is precisely to imply that it is a thing of such a sort. But something which is *in*capable of being named by a name 'for' a cat will be thus incapable just because it is *not* a thing of such a sort, and consequently it cannot truly be said that such a thing *is a cat* in the sense just used. Accordingly, there can as far as I can see be no univocal sense of 'being predicatively a cat' which is applicable equally to things which are and to things which are not capable of being named by a name 'for' a cat: which is precisely why a distinctive 'is' of constitution must be recognised, as I explained in my earlier article. If Noonan would have us take seriously the thesis he entertains, it is at the very least incumbent upon him to explain what *he* means (a) by something 'being predicatively a cat' and (b) by something being 'capable' or 'incapable' of being named by a name *for* a cat.

University of Durham,
50 Old Elvet,
Durham, DH1 3HN

©E. J. Lowe 1986

THE AVOIDANCE OF THE PROBLEM OF EVIL: A REPLY TO McGRATH

By ROGER CRISP

IN a recent article, 'Evil and the Existence of a Finite God' (ANALYSIS 46.1, January 1986), P. J. McGrath claims that a theist cannot avoid the problem of evil by limiting either the power or the goodness of God. I believe that both options remain open.

If we conceive of God as infinitely good, but not omnipotent, McGrath argues that such a deity is so weak as not to be worthy of our worship. Smallpox is an evil which no longer exists because humanity has wiped it out. If God is infinitely good, this implies that humanity is stronger than God in at least one area. McGrath has ignored, however, a thesis central in many theistic traditions that there are *two* powers at work in the universe, one infinitely good, the other infinitely evil. Thus, it could be that God is extremely powerful, and it is only with his help that humanity can overcome great evils such as smallpox. Such a deity is far more powerful than humanity, and, if we believe that there is more good than evil in the universe, the most powerful being in the universe. A being of this character would appear to be a quite proper object of worship.

If we adopt the second option, and conceive of God as omnipotent, but limited in goodness, McGrath claims that a God who allows the sort of evils which exist in the world would be more akin to a demon than to a deity, and thus unworthy of worship. Here, McGrath fails to notice another thesis important in a number of theistic traditions: the so-called 'Acts and Omissions Doctrine'. It may be that God is not as culpable for allowing certain evils to occur – through, say, the actions of a weaker but infinitely evil power, or human sin, or disasters of natural origin – as he would be had he acted to bring those evils about. Thus, if God only acts to bring about goods, if some of these goods are significant, and if God is not as responsible for his omissions as for his acts, he would appear at least to be worthy of worship, if not as worthy as he would be were he to act so as to prevent the evils in the world.

A third option suggested by McGrath is to conceive of God as limited in both goodness and power. If my arguments are correct, it is not necessary for a theist to retreat on both fronts, although I would claim that the arguments suggest that even such a God may sensibly be considered divine, and hence deserving of our worship.

St Anne's College,
Oxford OX2 6HS

© ROGER CRISP 1986

PLATO

Jobs for Philosophers

PLATO (Philosophy Lecturing And Teaching Opportunities) is a charitable scheme recently set up by philosophers from Bristol and Manchester Universities. Its objectives are, in general, to represent the interests of qualified British philosophers who are working in circumstances of economic hardship or insecurity. A more specific intention is to raise sufficient funds to provide for a limited number of academic fellowships, tenable (in principle) at the institution of the candidate's choice, but initially at the University of Bristol.

The organisers would welcome the advice and support of philosophers in all countries in pursuing these objectives.

Dean Buckner
University of Bristol
Department of Philosophy
9 Woodland Road
Bristol BS8 1TB

David Lamb
Department of Philosophy
The University
Manchester M13 9PL

The Monist *An International Quarterly Journal of General Philosophical Inquiry*

Editor, John HOSPERS

Editorial Board: William P. Alston, Monroe C. Beardsley, Dagfin Follesdal, R. Barcan Marcus, Joseph Margolis, Richard Martin, Wallace I. Matson, Joseph Owens, Richard Rorty, Wilfrid Sellars, Anita Silvers, John E. Smith, Richard Wasserstrom.

Managing Editor: Sherwood J. B. SUGDEN

EACH ISSUE IS LIMITED TO ARTICLES ON A SINGLE GENERAL TOPIC. Communicate in advance with the Editor for special instructions. Two nonreturnable copies (one copy outside the U.S. and Canada) should be submitted NINE MONTHS PRIOR to the announced publiation date. Papers should be prepared for anonymous submission to readers, and should be from 10 to 20 double-spaced typewritten pages in length, including notes.

SCHEDULED PUBLICATION DATES:

Vol. 70, No. 1	Jan.,	1987	Animal Rights
Vol. 70, No. 2	Apr.,	1987	Irrationality
Vol. 70, No. 3	July,	1987	The Ethics of Nuclear Warfare
Vol. 70, No. 4	Oct.,	1987	Thomas Reid and His Contemporaries
Vol. 71, No. 1	Jan.,	1988	Philosophical Problems of Space Exploration
Vol. 71, No. 2	Apr.,	1988	Aesthetics and the Histories of the Arts
Vol. 71, No. 3	July,	1988	Justification
Vol. 71, No. 4	Oct.,	1988	Descartes and His Contemporaries

Editorial Office: School of Philosophy, University of Southern California
Los Angeles, CA 90089
Business Office: Box 600, La Salle, Illinois 61301

SUBSCRIPTION RATES: United States: Annual (4 issues): Institutions, $24.00; Individuals, $15.00. Single copies: $6.00. Foreign postage: Add 50 cents to single copy rate or $1.00 to subscription.

INQUIRY *Editor: Alastair Hannay*

Vol. 29 (1986), No. 1

Symposium on Jon Elster's *Making Sense of Marx*	
Elster's Marx	Michael Taylor
Historical Materialism and Functional Explanation	Allen Wood
Making Nonsense of Marx	Scott Meikle
Making Sense of Elster	Cliff Slaughter
Is it Worth Making Sense of Marx?	Douglass C. North
Replies	Jon Elster
The Fate of the Enlightenment: Reply to Kekes	Nicholas Maxwell
Trivializing Knowledge: A Review of Popper's *Postscript*	Paul Feyerabend
Psychoanalysis, Science or Insight? (Adolf Grünbaum: *The Foundations of Psychoanalysis: A Philosophical Critique)*	Robert A. Sharpe

Subscription price 1986:
Institutions: USD 53.00 Individuals: USD 27.00

Published quarterly by Norwegian University Press, P.O. Box 2959 Tøyen, 0608 Oslo 6, Norway — P.O. Box 258, Irvington-on-Hudson, New York 10533, USA. Editorial address: Institute of Philosophy, P.O. Box 1024 Blindern, 0315 Oslo 3, Norway.

NOTES

The ANALYSIS Committee consists of: Chairman, P. T. Geach; Secretary, T. R. Baldwin; T. S. Champlin, M. Clark, J. P. Dancy, Andrew Harrison, B. J. Heal, R. F. Holland, J. H. McDowell, D. H. Mellor, H. W. Noonan, P. Smith, M. A. Stewart, R. G. Swinburne, A. R. White, C. J. F. Williams, Peter Winch. This committee is responsible for appointing and advising the Editor and for the general policy of the paper.

SUBSCRIPTIONS. ANALYSIS is published by Basil Blackwell Ltd, 108 Cowley Road, Oxford OX4 1JF, England and P.O. Box 1320, Murray Hill Station, NY 10156, U.S.A. Each volume comprises four numbers, each of at least 48 pages, appearing in January, March, June and October. Orders should be sent to the publisher. Subscriptions prices for 1986 are as follows:

	U.K.	Overseas (incl. Eire)	U.S.A. & Canada
Institutions	£11.75	£14.50	U.S. $33.00
Individuals	£7.85	£9.75	U.S. $21.00

BACK ISSUES. Volumes 1–37 are now available in complete volumes in an offset reprint from Swetz & Zeitlinger, 471 Keizersgracht, Amsterdam, Holland. Recent back issues are available from the publisher at the current single issue price.

MICROFILM copies of complete volumes are available from Volume 1 onwards from University Microfilms International Ltd, 30–32 Mortimer Street, London W1N 7RA, England, or 300 North Zeeb Road, Ann Arbor, Michigan 48106, U.S.A.

ADVERTISING. For details contact Bill Drake, telephone (0869) 38477, or write c/o the publishers.

CONTRIBUTIONS. Articles submitted for publication should be addressed to Mr Christopher Kirwan, Exeter College, Oxford OX1 3DP.

Contributors are asked to note the following requirements. Articles should normally not exceed 3,000 words in length, including references and footnotes. They should be type written in double spacing on one side of the paper only; photographic copies are acceptable. Footnotes should be kept to a minimum and wherever possible avoided altogether. Inset quotations and formulae should be well spaced from the text preceding and following them. The spacing of symbols within formulae should be carefully attended to. Single quotation marks should be used, except for purposes of internal quotation and "scare" quotes. Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

Typescripts of rejected articles will be returned only if self-addressed envelopes with correct postage are provided (in the case of overseas contributors, international reply coupons sufficient for airmail).

Proofs of accepted articles will be sent to authors for correction, together with information about free copies. Typescripts of accepted articles will be retained by the Editor on the assumption that authors have kept their own copies.

Apart from correction of printers' errors, only very minor changes in proof will normally be accepted by the Editor, to avoid delay and expense in production.

Typeset by Unicus Graphics Ltd, Horsham, Printed and bound in Great Britain by Whitstable Litho Ltd, Whitstable

Vol. 46 No. 4 (New Series No. 212) October 198

ANALYSIS

Edited by

CHRISTOPHER KIRWAN

CONTENTS

Open letter to the international philosophical community	PETER ALEXANDER *et al.*
Tense and truth conditions	GRAHAM PRIEST
Tense's tenseless truth conditions	D. H. MELLOR
How significant is the use/mention distinction?	A. W. MOORE
Indeterminacy of identity	MONTE COOK
On sets and worlds: a reply to Menzel	PATRICK GRIM
False stipulation and semantical paradox	LAURENCE GOLDSTEIN
On a supposed temporal/modal parallel	E. J. LOWE
Is there a temporal slippery slope paradox?	D. E. OVER
Reply to Over	E. J. LOWE
Reply to Lowe	D. E. OVER
Tarski – an ambiguity	RICHARD C. JENNINGS
Reply to Garrett	HAROLD W. NOONAN
Grayling, Feyerabend and the constancy of sense	JOHN SHAND
The possibility of transparent white	DAVID H. SANFORD
Noonan on temporal parts	GRAHAM SPINKS
More paradoxical epistemics	MARTIN HOLLIS
Reply to Lowe	HAROLD W. NOONAN
Superman, the image	ANDY CLARK
Correction	

BASIL BLACKWELL PUBLISHER · OXFORD

ISSN 0003-2638

BLACKWELL

Pragmatism without Foundations

Reconciling Realism and Relativism

JOSEPH MARGOLIS

Addressing the central question of how traditionalist claims to truth stand up in the light of recent pragmatist and relativist assumptions, Joseph Margolis's elegant and controversial analysis sets out to defend a novel version of relativism against the usual charges that it must be incoherent or self-defeating.

352 pages, **£25.00** (0 631 15034 X)

For Truth in Semantics

ANTHONY APPIAH

Anthony Appiah provides a fresh approach to the debate between realism and antirealism, arguing that antirealism fails in its attack on truth-based semantic theories, and that many of its positive claims are in fact false. His book adds a new dimension to the debate, and will open up new aspects of the subject.

Philosophical Theory

208 pages, **£22.50** (0 631 14596 6)

The Dynamics of Belief

A Normative Logic

PETER FORREST

In *The Dynamics of Belief* Peter Forrest argues strongly for the importance of logic as a normative, or prescriptive, theory. Taking a radical yet informed approach, he does not merely assess the rival claims of the various logics available, but considers what principles *ought* to be followed in the developments of belief.

Philosophical Theory

224 pages, **£22.50** (0 631 14619 9)

Objects of All Sorts

A Philosophical Grammar

VINCENT DESCOMBES

In this brilliant and engaging book, Vincent Descombes examines the nature of objects in contemporary thought. Originally published as *Grammaire d'objets en tous genres* in 1983, and now appearing with a new preface by the author, *Objects of All Sorts* applies the methods of analytical philosophy to a broad range of critical and philosophical issues.

256 pages, **£25.00** (0 631 15081 1)

Dialectics of Labour

Marx and his Relation to Hegel

C. J. ARTHUR

The first full-length, detailed study of Marx's 'Economic and Philosophical Manuscripts of 1844', this book outlines Marx's theory of alienation and considers his relations to Hegel. C. J. Arthur assesses the importance and limitations of the 1844 manuscripts, and investigates their links with Marx's later works and his hopes for socialism.

192 pages, **£22.50** (0 631 15218 0)

Basil Blackwell

108 Cowley Road, Oxford OX4 1JF

Suite 1503, 432 Park Avenue South, New York NY 10016

Law and Philosophy

An International Journal for Jurisprudence and Legal Philosophy

Editor

ALAN MABE, *Dept. of Philosophy, Florida State University, Tallahassee, U.S.A.*

Law and Philosophy is a forum for work in law and philosophy which is of common interest to members of the two disciplines of Jurisprudence and Legal Philosphy. It is open to all approaches in both fields and to work in any of the major legal traditions – common law, civil law, or the socialist tradition. The editors encourage papers which exhibit philosophical reflection on the law informed by a knowledge of the law, and legal analysis informed by philosophical methods and principles. Among others, the following topics are dealt with: theories of law, justice and law, morality and the law, legal validity, the nature of law and legal systems, punishment and responsibility, adjudication, legal reasoning, obligation and disobedience, professional ethics, epistemological issues involved in evidence and procedure, philosophical investigations of conceptual and foundational questions in law and legal systems, the impact of legal developments on philosophy, the relations of law and legal institutions to other social institutions, the relation of legal and political theory, the economic dimension of issues in jurisprudence and legal philosophy, the nature and development of legal concepts, and exploration of problems in particular areas of the law.

Subscription Information ISSN 0167–5249
1986, Vol. 5 (3 issues)
Institutional rate: Dfl. 201.00/US $72.00 incl. postage/handling
Private rate: Dfl. 80.00/US $28.00 incl. postage/handling

Private subscriptions should be sent direct to the publishers.

D. REIDEL PUBLISHING COMPANY

A Member of the Kluwer Academic Publishers Group

P.O. Box 17, 3300 AA Dordrecht, The Netherlands
190 Old Derby Street, Hingham MA 02043, U.S.A.
Falcon House, Queen Square, Lancaster LA1 1RN, U.K.

OPEN LETTER TO THE INTERNATIONAL PHILOSOPHICAL COMMUNITY

WE view with alarm the effect on philosophy in the United Kingdom of recent and prospective reductions in government funding to higher education.

Philosophy is among the splendours of the British intellectual tradition, which for centuries has nourished and enriched other disciplines. The continuance of this country's eminence in the subject can be ensured only by the steady recruitment of young people as professional philosophers.

We are glad to encourage and support the activities of PLATO in the pursuit of its stated objectives.*

Professor Peter Alexander
Professor Sir Alfred Ayer
Professor Sir Isaiah Berlin
Dr Simon Blackburn
Professor Myles Burnyeat
Professor F. C. Copleston
Professor Robert Cross
Professor Michael Dummett
Professor J. N. Findlay
Philippa Foot
Franks
P. L. Gardiner
Professor Ernest Gellner
Jonathan Glover
Martin Hollis
Professor Ted Honderich
Dr A. J. P. Kenny
Christopher Kirwan
Professor Stephan Körner
Professor H. D. Lewis
Dr Oonagh McDonald M.P.
Bryan Magee
Professor D. H. Mellor
Mary Midgley
Alan Montefiore
Iris Murdoch
Professor David Pears
Professor Roger Scruton
Professor Quentin Skinner
Professor Richard Sorabji
Professor Sir Peter Strawson
Professor G. N. Vesey
Sir Geoffrey Warnock
Professor R. A. Wollheim
Professor Crispin Wright

---

* PLATO (Philosophy lecturing and Teaching Opportunities, c/o Department of Philosophy, 9 Woodland Road, Bristol BS8 1TB) has the objectives of representing and protecting the interests of British philosophers who are working in circumstances of economic hardship or insecurity.

TENSE AND TRUTH CONDITIONS

By GRAHAM PRIEST

I INTRODUCTION

IS tense, that is, pastness, presentness and futurity, an objective feature of reality, or is it merely a subjective 'epiphenomenon'. This thorny subject has been the locus of a great deal of philosophical debate. A notable and recent instalment in the debate is Hugh Mellor's book *Real Time* (Cambridge U.P., 1984; all page references are to this). The book contains discussion of many issues, but the heart of it is the argument, with its dénouement in chapter 8, that tense is unreal. The chapter starts by using McTaggart's famous argument to try to show this. I do not think the discussion here adds significantly to the topic, and I doubt whether those unpersuaded by McTaggart will be persuaded by Mellor (in particular, no A-theorist who is awake will accede to the claim, p. 93 (2), that every event is past, present and future when the 'is' is significantly tensed). The second part of the chapter (pp. 98–102) gives a somewhat different argument for the same conclusion. According to Mellor this is just a modern re-working of McTaggart. I confess that this is less than clear to me, and I suspect that Mellor is unduly modest. At any rate, the argument can be formulated and discussed without any reference to McTaggart. It is this argument I wish to discuss and, I hope, show to be wrong. I will put the argument in my own terms without, I trust, doing Mellor any injustice.

II THE ARGUMENT

Is tense real? Like many philosophical issues, much of the problem is getting the question in a definite enough form to be able to say something intelligent about it. Mellor glosses the question (in a way of which I thoroughly approve) as follows. Consider the set of true sentences (of, say, English). What makes these true are the facts. We do not have to reify facts as traditional correspondence theorists did. The facts are just whatever make claims true if they are true. Reality is constituted by the facts. Hence, if we can account for the truth of all true sentences without appealing to a certain notion, such is not a part of reality. We say what makes a sentence true, if it is, by giving its truth conditions. Thus, if we can spell out the truth conditions of all sentences (true and false) without mentioning a notion, that notion is certainly not part of reality. A sufficient condition for something's being unreal is, therefore, that it is irrelevant to the truth conditions of all sentences. Applying this:

tense is unreal if we can give the truth conditions of all sentences without using tensed sentences. (It is, I suppose, possible to object to this test for reality that there might be parts of reality that are not the truth makers for any sentences, and thus are ineffable. We need not enter this debate here. For tense is certainly quite effable.)

The next question is whether we can give the truth conditions of all sentences of, say, English without invoking the notion of tense. To this Mellor gives the answer 'yes'. There are two relevant cases to distinguish here depending on whether the verb of the sentence is significantly tensed or not. It will be absolutely crucial that we distinguish between tensed and untensed verbs, especially the copula. I will therefore henceforth italicize all and only tensed verbs. Thus I will write, e.g., 'The Battle of Waterloo *occurred* in 1815', but 'The Battle of Waterloo occurs in 1815' where both sentences are thought of as stating truths. Now, to give the truth conditions of untensed sentences without employing tensed sentences is simplicity itself. The homophonic T-scheme will do. Thus, for example:

(1) 'The Battle of Waterloo occurs in 1815' is true iff the Battle of Waterloo occurs in 1815.

For tensed sentences we have to be a little more cunning. Mellor employs a well known device. It now becomes important to decide of what sort of thing truth is to be predicated. Mellor elects the class of utterances (or better, utterance-event-tokens). The temporal location of an utterance can then be used in the relevant truth conditions. Thus, suppose that *u* is an utterance of 'The Battle of Waterloo *occurred* in 1815'. The appropriate instance of the T-scheme is:

(2) *u* is true iff the Battle of Waterloo occurs in 1815 and 1815 occurs before *u*.

A similar dodge can obviously be performed with present and future tenses, and, as Mellor notes (p. 42ff), with compound tenses.

We have seen that the truth conditions of all sentences, tensed and untensed, can be given without using tensed sentences at all. Tense, therefore, is not a part of reality. As Mellor puts it (p. 102):

> The sole function of tensed facts is to make tensed sentences and judgements true or false. But that job is already done by the tenseless facts that fix the truth-values of all tensed sentence and judgement tokens. Provided a token of '*e* is past' is later than *e*, it is true. Nothing else about *e* and it matters a jot: in particular, no tensed fact about them matters. It is immaterial, for a start, where *e* and the token are in the A series; and if that is not material, no more *recherché* tensed fact can be. Similarly for tokens of all other tensed types. Their tenseless truth conditions leave tensed facts no scope for determining their truth values. But these facts by definition determine their truth values. So in reality there are no such facts.

The argument against tense does not, however, have to depend on an appeal to Ockham's razor. For the supposition that there are tensed facts, that is, that tensed notions are to be used to give truth conditions, leads to contradictions. Thus suppose we try to use them to give the truth conditions of tensed sentences, obviously the most favourable case. Let *u* be any utterance of 'The Battle of Waterloo *occurred* in 1815.' Then the appropriate instance of the T-scheme would be:

(4) *u* is true iff the Battle of Waterloo *occurred* in 1815.

Then, applying (4) and using our knowledge of history, we can infer that *u* is true. However, (4) was equally applicable in 1800 (assuming that reference can be made to future events – an assumption that we need not here challenge). Moreover, since then the Battle of Waterloo *had not happened*, *u* is false. Thus *u* is both true and false. This contradiction is, according to Mellor, essentially the same as that which eventuates in McTaggart's argument. Whether or not that is so, it is a contradiction, and therefore the assumption that there are tensed facts is incorrect. To quote Mellor again (p. 100):

> ... giving any tensed sentence non-token-reflexive truth conditions, tensed or tenseless, always leads to contradiction. ... (such truth conditions) would give all tokens of the same type the same truth-value, regardless of their date, thus inevitably contradicting the truth value of some tokens of the tensed sentence, which differ from date to date. ... Since the truth-value of tensed tokens is never independent of their A-series position, giving them now all the same truth value will inevitably make some past or future tokens both true and false.

III THE REPLY

The argument is wrong. To explain why it is wrong, I will start by giving a parallel argument. To see what this is, note first that in the truth predicate used in Mellor's argument the 'is' of 'is true' is *not* significantly tensed. (I have, according to my convention, not, therefore, italicized it.) If this is not clear, merely note that in the argument of the last paragraph, if the possession of truth were significantly tensed, all that we could conclude is that *u is* true now but *was* false in 1800. It would therefore follow that a numerically unique utterance can change its truth value over time. But this is no more problematic than its changing its colour (say, because of fading). Now, suppose that instead we take the truth predicate to be significantly tensed. Then we may argue exactly as Mellor does, except, this time...

We can prove that *tenselessness* is not a part of reality. We do this by showing that the truth conditions of all sentences can be given in tensed terms. How to do this for tensed sentences is entirely obvious. The homophonic T-scheme will do. Thus, for example:

(5) 'The Battle of Waterloo *occurred* in 1815' *is* true iff the Battle of Waterloo *occurred* in 1815.

The truth value of the utterance will change over time: the utterance, whenever it is made, *is* true in 1986 and *was* false in 1800, since then the Battle of Waterloo *had not happened*; but as I noted, this is unproblematic.

To state the truth conditions of untensed sentences we have to employ a little more cunning – but not much. Let us define, for any verb 'to *v*' another verb 'to eternally *v*' thus: '*x eternally v's*' is '*x has v'ed* or *x v's* or *x will v*'. The important thing about a sentence whose verb is an eternal one is, as the name implies, that if it is true at any time it is true at all times. The truth conditions of untensed sentences can now be given in eternal terms. Thus for example:

'The Battle of Waterloo occurs in 1815' *is* true iff the Battle of Waterloo *eternally occurs* in 1815.

This is perfectly correct. Though both sides of the biconditional are significantly tensed, both sides have the same truth value, namely true, at all times. Since the truth conditions of all sentences can be given in tensed terms, tenselessness is not a part of reality.

So far this argument, with its diametrically opposed conclusion, would seem as good, or as bad, as Mellor's. But Mellor's argument did not end there. The supposition that there are tensed facts actually leads to contradiction. But now that we have significantly tensed the truth predicate, the supposition that there are untensed facts leads to contradiction. Suppose, for example, that *u* is an utterance (at any time) of the sentence 'The Battle of Waterloo *occurred* in 1815', and we try to use an untensed sentence to give its truth condition thus:

(6) *u is* true iff the Battle of Waterloo occurs in 1815

Then applying this now, we infer that *u is* true. However, the truth conditions were equally applicable in 1800, and applying them then, it follows that *u was* then true. But this contradicts the plain fact that it *was* not. (Imagine that *u* was uttered in 1800.) Nor will it help to give token reflexive truth-conditions *à la* Mellor. For then the truth conditions are:

(7) *u is* true iff the Battle of Waterloo occurs in 1815 and 1815 occurs before *u*.

Applying (7) now we infer that *u is* true; but, as before, (7) was equally applicable in 1800, and applying it then we conclude that *u was* true in 1800, which contradicts the fact that it *was* not.

The problems that the rival truth conditions run into are, in fact, nothing to do with the reality, or otherwise, of anything, but merely reflect the fact that whether one takes truth to be tensed or untensed it is possible for an imbalance to arise in the T-scheme. If one can produce an instance of the T-scheme that has one side significantly (and not eternally) tensed, while the other side is not, then applying it at times other than the present will produce trouble. In the case of Mellor's argument the left-hand side of (4) is not tensed whilst the right-hand side is. In the case of the dual argument, the left-hand sides of (6) and (7) are significantly tensed (because of the tensed truth predicate) whilst their right-hand sides are not.

We see then that Mellor's argument against the reality of tense does not work; the conclusion is merely the result of a silent assumption that truth is untensed. The tensedness of truth is, of course, exactly what someone who takes tense to be part of reality is likely to claim. It would appear, then, that Mellor's argument tacitly begs the question.

IV CONCLUSIONS

The next obvious question is whether one can give an independent argument for the claim that the notion of truth used in giving truth conditions must be untensed. Logicians, it is true, usually talk in tenseless terms. For example the papers comprising the burgeoned literature on truth conditions would seem to make this assumption. (And this maybe explains why Mellor makes the assumption without apparent question.) This would seem, however, to be an accident. For as we have just seen, one can give truth conditions quite adequately whether truth is taken to be tensed or to be untensed. The approaches are not logically equivalent. On the tenseless approach utterances have a truth value fixed for all time; and if the sentence is tensed, this truth value will be determined in part by the date of the utterance. On the tensed approach a numerically unique utterance – even one which exists for only a moment – may change its truth value over time; and the date of utterance will be irrelevant to this truth value (unless, of course, the utterance is to the effect that it does or does not occur at a certain date). However, there seems to me to be no particular reason for preferring one of these approaches to the other. You can think of things in either way provided that you are clear about how it is that you are thinking about them, and do not, of course, equivocate. Maybe this is wrong and there are decisive reasons for adopting one approach rather than the other. However, unless and until this is shown the argument from truth conditions goes nowhere.

University of Western Australia,
Nedlands 6009, Australia

© GRAHAM PRIEST 1986

TENSE'S TENSELESS TRUTH CONDITIONS

By D. H. MELLOR

I ARGUE in *Real Time* that reality is tenseless because all tensed truths have tenseless truth conditions. In his 'Tense and truth conditions' (ANALYSIS 46.4, October 1986, pp. 162–6), Graham Priest tries to rebut my argument by giving a parallel one to show that all tenseless truths have tensed truth conditions. His parallel fails, because to draw it he has to take what I call 'tense' to be a feature of verbs, despite initially defining it as 'pastness, presentness and futurity' and saying rightly that the question at issue is whether tense so defined is 'an objective feature of reality' (p. 162). This of course is a question about time, not about verbs; but as Priest is still by no means alone in confounding the two, I must first redistinguish them; and to mark the distinction I will adapt a device of Priest's by using italics for *verbal tense* and capitals for PAST, PRESENT and FUTURE.

What I call TENSES are temporal positions in McTaggart's *A* series, 'that series of positions which runs from the far past through the near past to the present, and then from the present through the near future to the far future' (McTaggart 1927, p. 10). TENSES thus include not only the rather extended positions PAST and FUTURE, but also more restricted PAST or FUTURE positions like YESTERDAY and NEXT YEAR, and PRESENT ones like TODAY and THIS YEAR.

I then call a thought, statement or sentence TENSED if it explicitly or implicitly ascribes a TENSE to something, its TENSE being the TENSE it ascribes. Now English grammar usually requires TENSED sentences to use *tensed* verbs, i.e. inflected verbs whose inflection depends on whether the TENSE ascribed is PAST, PRESENT or FUTURE. But not all languages use *tensed* verbs, and even in English they do not suffice to ascribe more restricted TENSES, for which we need adverbs or phrases like 'yesterday', 'today' and 'next year': and these make *tensed* verbs redundant. The TENSE of 'John died yesterday', for instance, is fixed by the meaning of 'yesterday', and since YESTERDAY is a PAST TENSE, English grammar makes us inflect the verb accordingly. But the sentence would convey the same temporal information even if its verb were not inflected: i.e. even if in English we said (e.g.) 'John dies' whether we are going on to add 'yesterday', 'today' or 'tomorrow'.

The worst excesses of the linguistic revolution in philosophy may be over, but it still needs stressing that the concept of TENSE in no way depends on verbal inflection. Even the simple PAST, PRESENT and FUTURE cannot be defined as the temporal reference of corresponding *tenses*, if only because these TENSES and *tenses* often do *not* correspond. (Think for instance of 'She now leaves next week', 'She will be leaving Heathrow at this very minute' and 'If only she left tomorrow . . . '.) The fact is that *tense* in English is a function of TENSE (and other factors) but not vice versa, so that mastery of it demands a *prior* grasp of TENSE. And the correspondence between TENSE and *tense* that we do have comes from rules like: when talking in simple declarative sentences about what is PAST, use verbal inflection '*V*-ed' (e.g. 'died'). In short, the *past tense* is named after the PAST TENSE, not the other way round (*cf.* Dudman 1983).

So what Priest calls 'eternal' verbs are not, as he claims, defined in terms of *tensed* ones (p. 165): they are just verbs, like my 'dies' above, which do not inflect with the TENSE a sentence is used to ascribe. The verb 'is' in

(*S*) John is dead

for instance, can perfectly well be read as such a verb. *S* will still be TENSED: PAST, of course, because what 'dead' makes it mean is that John's death is PAST.

But, it may be objected, the 'is' in *S* cannot be 'eternal', because John's death was not always PAST: *S* could not be truly asserted at all times. So it is never asserted as being true always, only as being true at the time of assertion, i.e. at what is then the PRESENT time. So – the objection runs – in order to show that John's death is only being said to be NOW PAST, the 'is' in *S* must be 'present' in the sense of having 'a *tense* . . . used when the action or event described is occurring at the time of utterance' (*Collins Dictionary of the English Language* 1982, p. 1157).

This is nonsense. Of course asserting any sentence *X* at any time is asserting it as being true then. But that cannot – on pain of an obvious vicious regress – mean asserting '*X is* [NOW] true'. Nor can it mean that the main verb in *X* must be *present tense* in the above sense, or we could not assert sentences whose main verbs are *past* or *future tense*. All it means is that asserting *X* by producing a token of it is claiming truth for that token, and what follows for tokens produced at other times then depends on *X* (for TENSELESS *X*, that they would all be true; for *S*, being PAST TENSE, only that later ones would be). But this is quite enough to make the PAST TENSE ascribed by a token of *S* relative to when the token is produced: *S* need not also contain a *present tense* 'is'. *S* would have just the same PAST TENSE if its 'is' were as 'eternal' and semantically dispensable as the 'is' in '3 is odd'. So who is to say it isn't? The 'is' in 'John is dead' is in fact '*present*' only in the *Collins Dictionary's* innocuous alternative sense of a *tense* used 'when the speaker does not wish to make any explicit temporal reference' (*loc. cit.*).

But how about 'is true' in '*S* is true'? Priest says the conclusion of my argument 'is merely the result of a silent assumption that

TENSE'S TENSELESS TRUTH CONDITIONS

By D. H. MELLOR

I ARGUE in *Real Time* that reality is tenseless because all tensed truths have tenseless truth conditions. In his 'Tense and truth conditions' (ANALYSIS 46.4, October 1986, pp. 162–6), Graham Priest tries to rebut my argument by giving a parallel one to show that all tenseless truths have tensed truth conditions. His parallel fails, because to draw it he has to take what I call 'tense' to be a feature of verbs, despite initially defining it as 'pastness, presentness and futurity' and saying rightly that the question at issue is whether tense so defined is 'an objective feature of reality' (p. 162). This of course is a question about time, not about verbs; but as Priest is still by no means alone in confounding the two, I must first redistinguish them; and to mark the distinction I will adapt a device of Priest's by using italics for *verbal tense* and capitals for PAST, PRESENT and FUTURE.

What I call TENSES are temporal positions in McTaggart's *A* series, 'that series of positions which runs from the far past through the near past to the present, and then from the present through the near future to the far future' (McTaggart 1927, p. 10). TENSES thus include not only the rather extended positions PAST and FUTURE, but also more restricted PAST or FUTURE positions like YESTERDAY and NEXT YEAR, and PRESENT ones like TODAY and THIS YEAR.

I then call a thought, statement or sentence TENSED if it explicitly or implicitly ascribes a TENSE to something, its TENSE being the TENSE it ascribes. Now English grammar usually requires TENSED sentences to use *tensed* verbs, i.e. inflected verbs whose inflection depends on whether the TENSE ascribed is PAST, PRESENT or FUTURE. But not all languages use *tensed* verbs, and even in English they do not suffice to ascribe more restricted TENSES, for which we need adverbs or phrases like 'yesterday', 'today' and 'next year': and these make *tensed* verbs redundant. The TENSE of 'John died yesterday', for instance, is fixed by the meaning of 'yesterday', and since YESTERDAY is a PAST TENSE, English grammar makes us inflect the verb accordingly. But the sentence would convey the same temporal information even if its verb were not inflected: i.e. even if in English we said (e.g.) 'John dies' whether we are going on to add 'yesterday', 'today' or 'tomorrow'.

The worst excesses of the linguistic revolution in philosophy may be over, but it still needs stressing that the concept of TENSE in no way depends on verbal inflection. Even the simple PAST, PRESENT and FUTURE cannot be defined as the temporal reference of corresponding *tenses*, if only because these TENSES and *tenses* often do *not* correspond. (Think for instance of 'She now leaves next week', 'She will be leaving Heathrow at this very minute' and 'If only she left tomorrow . . . '.) The fact is that *tense* in English is a function of TENSE (and other factors) but not vice versa, so that mastery of it demands a *prior* grasp of TENSE. And the correspondence between TENSE and *tense* that we do have comes from rules like: when talking in simple declarative sentences about what is PAST, use verbal inflection '*V*-ed' (e.g. 'died'). In short, the *past tense* is named after the PAST TENSE, not the other way round (*cf.* Dudman 1983).

So what Priest calls 'eternal' verbs are not, as he claims, defined in terms of *tensed* ones (p. 165): they are just verbs, like my 'dies' above, which do not inflect with the TENSE a sentence is used to ascribe. The verb 'is' in

(*S*) John is dead

for instance, can perfectly well be read as such a verb. *S* will still be TENSED: PAST, of course, because what 'dead' makes it mean is that John's death is PAST.

But, it may be objected, the 'is' in *S* cannot be 'eternal', because John's death was not always PAST: *S* could not be truly asserted at all times. So it is never asserted as being true always, only as being true at the time of assertion, i.e. at what is then the PRESENT time. So – the objection runs – in order to show that John's death is only being said to be NOW PAST, the 'is' in *S* must be 'present' in the sense of having 'a *tense* . . . used when the action or event described is occurring at the time of utterance' (*Collins Dictionary of the English Language* 1982, p. 1157).

This is nonsense. Of course asserting any sentence *X* at any time is asserting it as being true then. But that cannot – on pain of an obvious vicious regress – mean asserting '*X is* [NOW] true'. Nor can it mean that the main verb in *X* must be *present tense* in the above sense, or we could not assert sentences whose main verbs are *past* or *future tense*. All it means is that asserting *X* by producing a token of it is claiming truth for that token, and what follows for tokens produced at other times then depends on *X* (for TENSELESS *X*, that they would all be true; for *S*, being PAST TENSE, only that later ones would be). But this is quite enough to make the PAST TENSE ascribed by a token of *S* relative to when the token is produced: *S* need not also contain a *present tense* 'is'. *S* would have just the same PAST TENSE if its 'is' were as 'eternal' and semantically dispensable as the 'is' in '3 is odd'. So who is to say it isn't? The 'is' in 'John is dead' is in fact '*present*' only in the *Collins Dictionary's* innocuous alternative sense of a *tense* used 'when the speaker does not wish to make any explicit temporal reference' (*loc. cit.*).

But how about 'is true' in '*S* is true'? Priest says the conclusion of my argument 'is merely the result of a silent assumption that truth is untensed' (p. 165). Very well, let me break the silence: truth is untensed. That is, the predicate 'is true' is not TENSED – and the 'is' in it is only innocuously *present tense*.

Whether 'is true' is TENSED may depend on what it is predicated of. It fortunately doesn't matter whether it is predicated of sentences, propositions or propositional attitudes. So we need only consider sentences, and the argument will apply *pari passu* to the other truth-bearers. But it may matter whether it is predicated of sentence *types* or their *tokens*. So I consider these in turn.

But first we must say what makes a predicate TENSED. Since a sentence is TENSED when it ascribes a TENSE, I take a predicate to be TENSED when it makes sentences containing it TENSED (or affects their TENSE). That is the obvious sense in which 'is past' and 'is dead' are TENSED predicates. It 'is true' in this sense a TENSED predicate of sentence types?

Not on the face of it. No TENSED nor TENSELESS sentence *type X* has any *A* series position, and calling it true never ascribes one to it. But calling 'is true' TENSED might also mean that it makes the truth of '*X* is true' depend on the *A* series position of something *X* is about. But when '*X* is true' is TENSED in that sense, what makes it so is not that 'is true' is TENSED but that *X* is, since '*X* is true' always has the same TENSE as *X* itself. For consider a token *Xt* of any *X* asserted at any time *t* (which may be given TENSELESSLY or as a TENSE: it makes no odds). Then imagine *Xt* replaced by a token *XTt* of '*X* is true' (*XT*). *XTt* will be true if and only if *Xt* was: since, as we have seen, asserting *XT* at *t* is asserting *XT* as being true then, it is at least asserting *that X* is true at *t*, i.e. that *Xt* would be true. Thus the truth values of *Xt* and of *XTt* are always the same function of *t*: *X* and *XT* always have the same truth conditions. In short, the predicate 'is true' is no more responsible for the TENSE of a TENSED '*X* is true' than the 'is' in 'John is dead' is responsible for the TENSE of that sentence type.

But 'is true' might still be a TENSED predicate of sentence *tokens*, which (unlike types) do have *A* series positions on which their truth might depend. Suppose for instance John was said to be dead just before, and again just after, he died at time *t*. Call his death '*D*', and these tokens of *S* respectively '*St*—' and '*St*+'. *D*'s TENSE, i.e. its *A* series position, of course keeps changing, as it becomes more and more PAST. But so do the *A* series positions of *St*— and *St*+, and their temporal *relations* to *D* never change: *St*— is always MORE PAST or LESS FUTURE, *St*+ always LESS PAST or MORE FUTURE, than *D*. But these are just McTaggart's *B* series relations: *St*— and *St*+ are always respectively *earlier* and *later* than *D*. The question is, what do the truth values of these tokens depend on: (i) their unchanging *B* series relations to *D*, or (ii) *D*'s own changing TENSE? If (i), 'is true' will not be a TENSED predicate of these tokens; and if (ii), it will be.

(i) says that *S*'s meaning makes those of its tokens earlier than *D* always false, and those later than *D* always true. Their truth values differ from each other, but never change. In particular, as shown in Figure 1, *St*— will be false, and *St*+ true, both before and after John dies.

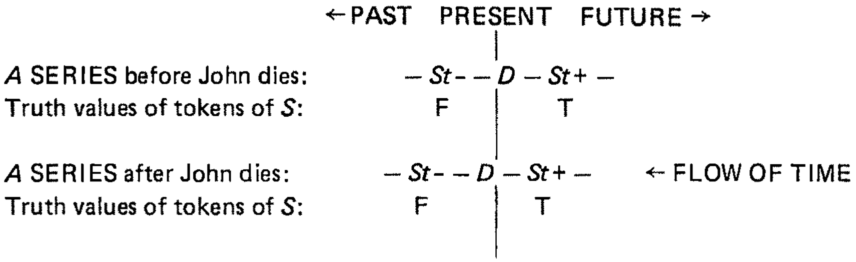

←PAST PRESENT FUTURE→
A SERIES before John dies:
– *St*– – *D* – *St*+ –
Truth values of tokens of *S*:
F
T
A SERIES after John dies:
– *St*– – *D* – *St*+ –
← FLOW OF TIME
Truth values of tokens of *S*:
F
T

Figure 1

(ii) on the other hand says that *S*'s meaning makes all its tokens false before John dies, and all true after. Their truth values change, but never differ from each other. In particular, as shown in Figure 2, *St*— and *St*+ will both be false before John dies, and both true afterwards.

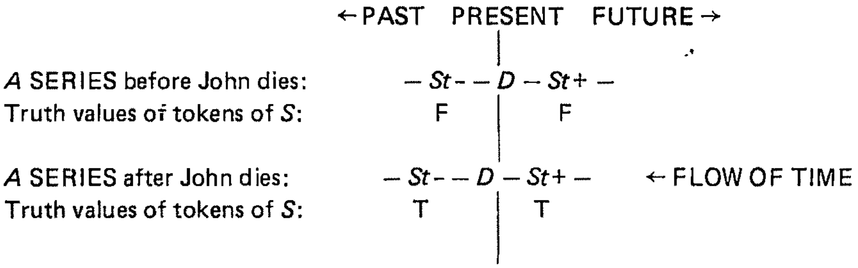

←PAST PRESENT FUTURE→
A SERIES before John dies:
– *St*– – *D* – *St*+ –
Truth values of tokens of *S*:
F
F
A SERIES after John dies:
– *St*– – *D* – *St*+ –
← FLOW OF TIME
Truth values of tokens of *S*:
T
T

Figure 2

But it is absurd to say, as (ii) does, that John's death posthumously verifies every premature announcement of it! It is like, in the personal and spatial analogues of TENSE, my impugning your utterance of 'I'm drunk' because *I*'m sober, or London's 'Cambridge 60' road signs because Cambridge isn't 60 miles from *here* (Cambridge). (*Cf.* Dummett 1973, pp. 385-6.) Those reactions would just show that I hadn't a clue what 'I'm drunk' and 'Cambridge 60' mean. And (ii) similarly grossly mistakes the meaning of 'John is dead'.

(i) gives *S* its right meaning with its token-reflexive *B* series truth conditions, and (*mutatis mutandis*) that of every other TENSED sentence. As Figure 1 shows, therefore, though some tokens of a TENSED sentence type will differ in truth value, their truth values never change: 'is true' is no more TENSED when predicated of TENSED sentence tokens than when predicated of TENSED sentence types.

But then TENSED sentence types cannot be made true by TENSED facts, like the fact that John's death is PAST. For suppose this alleged fact makes *S* true. Until John dies it is *not* a fact, so until then *S* is false, i.e. all its tokens are false. And afterwards, when it *is* a fact, *S* is true, i.e. all its tokens are true. That is, their truth values are as shown in Figure 2: never differing one from another, and all changing together from false to true when *D* becomes PAST.

But as we have just seen, that is not so: the truth values of *S*'s tokens never change, and those earlier than *D* always differ in truth value from those later than it. But this is flatly inconsistent with ever assigning any uniform but changeable truth value to all of them, i.e. with taking *S* itself to be made true by some TENSED fact. (This is what I take to underlie McTaggart's contradiction.) And as for *S*, so for all other TENSED sentences. Their token-reflexive *B* series truth conditions prevent them ever being made true by TENSED facts. But, as Priest (p. 162) and I agree, facts just *are* what make truth conditions obtain. So since no truth conditions can ever be made to obtain by TENSED facts, there can in reality be no such facts. In particular, there cannot really be the two *A* series depicted in Figure 1, only the single *B* series of Figure 3:

← EARLIER LATER →

B SERIES: – *St*– – *D* – *St*+ –
Truth values of tokens of *S*: F T

Figure 3

McTaggart of course notoriously argued that time needs *A* series change, so that *B* series relations must derive from a real *A* series in order to be temporal: 'later' must really mean MORE FUTURE or LESS PAST. That argument I rebut at length in *Real Time*, in particular by showing how time can be made the dimension of change, and our experience of time flowing can be accounted for, without any appeal to an *A* series. Here I will only repeat my answer to the myth (still current e.g. in McGinn 1983, pp. 104–6) that *A* series positions are given to us in perception, since 'we perceive events in time as being present' (McTaggart 1927, p. 11). This is just not true. We perceive nothing beyond our own experience as being PRESENT, any more than we perceive it as being PAST or FUTURE. What we see through telescopes no more *looks* millions of years PAST than it looks PRESENT (if it did, celestial chronology would be easier than it is); and claims to see the

FUTURE in a crystal ball are not refutable by any visible PASTNESS or PRESENCE in what it reveals.

Whereas we do often see that one event is *later* than another, wherever they may be in the *A* series: we do so for instance whenever we see motion (or any other change) occur one way rather than the other. And ascriptions of TENSE must then conform to these directly observable *B* series relations: 'MORE FUTURE or LESS PAST' means *later*, not the other way round.

TENSED modes of thought and language arise not from passive perception of a TENSED world, but from the incessant need of agents to take timely action in a TENSELESS one (see *Real Time*, ch. 5, Perry 1979). But this does not make TENSE 'merely a subjective "epiphenomenon" ' (Priest, p. 162). That is not the alternative to its being 'an objective feature of reality'. TENSE is not a *subjective* feature of reality: like truth, it is not a feature of reality at all, except in the harmless sense that reality contains TENSED thought and language, much of it true. Like truth, TENSE is a feature of thoughts and sentences, not of what they are about (except, harmlessly, when they are about the TENSE or truth of other thoughts or sentences). So one might indeed call both TENSE and truth 'epiphenomena' of thought and language, or of their possessors. But that does not make either of them 'merely subjective'. They are both basic and indispensable as concepts; and both are equally objective. Maybe only subjects can think or use a language, but that platitude is no more what makes a thought or sentence TENSED than it is what makes it true. TENSED thoughts and sentence tokens differ from TENSELESS ones only by their truth depending on *when* they occur, and there is nothing subjective about that. After John's death, the thought that he is dead is not only indispensable to his grieving friends, and different in its truth conditions from any TENSELESS thought: like the notice of his death in the paper, the thought is also as objectively true as anything can be. For what makes both thought and notice true is just the wholly objective TENSELESS fact that they occur after John dies.

Darwin College, Cambridge CB3 9EU © D. H. MELLOR 1986

REFERENCES

V. H. Dudman, 'Tense and time in English verb clusters of the primary pattern', *Australian Journal of Linguistics*, 3, (1983) pp. 25–44.

M. Dummett, *Frege: Philosophy of Language*, Duckworth 1973.

C. McGinn, *The Subjective View*, Clarendon Press, 1983.

J. McT. E. McTaggart, *The Nature of Existence*, Vol. 2, Cambridge University Press, 1927.

D. H. Mellor, *Real Time*, Cambridge University Press, 1981.

J. Perry, 'The problem of the essential indexical', *Nous*, 13, (1979), pp. 3–21.

G. Priest, 'Tense and truth conditions', *Analysis*, 46.4, October 1986, pp. 162–6.

HOW SIGNIFICANT IS THE USE/MENTION DISTINCTION?

By A. W. MOORE

I

LET us take for granted that there is a distinction to be drawn between using an expression and mentioning it. And let us say that an expression is *employed* when it is put to any kind of service. Then I shall assume each of the following: that using an expression (in the relevant quasi-technical sense) involves employing it in such a way as to exploit (one of) its meaning(s); and that mentioning an expression involves employing it together with certain other linguistic devices which are themselves being used to demonstrate (or "to point to") some aspect of this employment. These assumptions take their cue from Donald Davidson's paper [1] (cf. also [3], p. 21 and [5]). Davidson there says that quotation marks are used to refer to expressions by demonstrating particular tokens of them. I have deliberately said something much looser than this. For one thing, as Davidson acknowledges, quotation marks are not the only device which enable us to mention expressions. Also, for reasons which will emerge, I prefer not to commit myself on the question of just what is being demonstrated when such devices are used.

Note that it is perfectly possible to employ an expression without either using it or mentioning it. Consider, for example, saying 'Cheese' when having a photograph taken. More interestingly, it is possible for an employment of an expression to be both a use and a mention of it. For example, one would be both using and mentioning the quoted words if one asserted the following sentence:

(1) Quine describes the totality of our knowledge and beliefs as '. . . a man-made fabric which impinges on experience only along the edges.'

Again, one would be both using and mentioning 'Arthur' if one said

(2) Arthur was so named after an uncle.

There are many ways of mentioning an expression other than by using quotation marks or referring phrases such as 'so named'. One way has just been illustrated: displaying an expression on a separate line of print (here with a numerical label). Writing an expression in italics is another way. Employing an expression after such a phrase as 'the word' or 'the expression' is another very common way. For example, in order to say that 'transcendental' is an esoteric word, one can use the following sentence:

(3) The word transcendental is esoteric.

Another common way of mentioning an expression is simply to employ it as a singular term. One could just as well have used

(4) Transcendental is an esoteric word.

If the expression is one which would normally be used as a singular term, an ambiguity arises. But the ambiguity will usually be easy to resolve. Compare how 'Elizabeth' would be employed in typical uses of

(5) Elizabeth has arrived

and

(6) My name is Elizabeth.

No doubt there are other ways of mentioning expressions besides these. The little that I have said so far leaves it unsettled whether a whole range of uses also count as mentions – for example, uses of expressions in scare-quotes or uses of expressions in italics when they are being defined. (Consider the scare-quoted 'to point to' and the italicized 'employed' in the opening paragraph of this essay.) I have certainly not provided a comprehensive and determinate catalogue of the devices which can be used to mention expressions. Nor am I interested in doing so. I think that anybody attempting to devise such a catalogue would find that, if it could be done at all, then eventually more or less arbitrary stipulations would be called for. For my main contention is this: the kind of significance which the use/mention distinction is usually thought to possess must in fact, if the distinction is to be a clear-cut one, reside elsewhere.

II

It is usually thought that the significance of the use/mention distinction consists in its coincidence with another distinction: the distinction between employing an expression in order to draw attention to, or to talk about, some aspect of the world, as determined by the expression's meaning, and employing it in order to draw attention to, or to talk about, the expression itself. This latter distinction is significant because it is sometimes easy to confuse these ways of employing expressions – with serious consequences.

It is undeniable that invoking the use/mention distinction will often serve to combat such confusion. Imagine, for example, somebody's challenging scientific pretensions to have discovered that heat is mean kinetic energy on the grounds that the latter, unlike the former, is a piece of technical terminology used only by a small group of initiates. But notions like what one is employing an expression for, what one is drawing attention to, and what one is talking about, are grounded in the pragmatics of communication. How they apply is a messy, complex, indeterminate matter, heavily dependent on particular circumstances and varying in degree from one case to the next. (What is Dylan Thomas employing the following words for when he writes, 'the sloeblack, slow, black, crowblack, fishingboat-bobbing sea'?) If the use/mention distinction is to be clear-cut – if there is to be a range of easily recognizable and specifiable devices which register when an expression has been mentioned – then, even if the distinction has an interesting bearing on what expressions are employed for, it cannot be right to insist on the coincidence outlined above; and this in turn means that it cannot be right to credit the use/mention distinction with the significance which it is usually thought to have.

The sentences (7) to (12) below help to bear this out.

(7) The city where Kant lived was then Königsberg but it has now changed to Kaliningrad
(8) Christopher can never remember his four-times table; he thinks that three fours are sixteen
(9) This is sepia
(10) Do you know the woman who works there called Catherine?
(11) 'This is ridiculous,' she thought to herself
(12) In a certain sense of 'know', nobody knows anything

Think about typical uses of these sentences. Imagine (9), for example, being used while pointing to something in order to provide an ostensive definition, in response to the question: what does 'sepia' mean? (On one fairly radical view, whereby the meaning of any given expression is continually evolving, this would not be fundamentally different from a supposedly non-defining use of (9).) Now consider what the underlined expressions would be being employed for. This would certainly not depend in the requisite way on whether they were being used or mentioned – provided, that is, it were transparent whether they were being used or mentioned, the quotation marks in (11), for example, signifying the latter. For excepting (12), the cases where it would be most natural to say that the expressions were being employed to draw attention to, or to make some claim about, themselves ((7) to (9)), would be precisely the cases where they were not being mentioned. And in (12) the two underlined expressions would naturally be said to be operating in tandem, in such a way that little further significance (beyond the demands of grammar) could attach to the fact that one, but not the other, was being mentioned.

Of course, one could deny that it was transparent which expressions were being mentioned in these cases. After all, sentences can have a misleading surface grammar (for example, they can contain ellipsis), and this means that an expression can be mentioned surreptitiously – or only apparently. So one could insist on a coincidence between the use/mention distinction and a distinction based on what expressions are being employed to draw attention to or to talk about, even in the face of examples such as these. For instance, one could say that 'Königsberg' was being mentioned in the first example and that it only seems otherwise because of an ellipsis. (The city was then *called* Königsberg.) On the other hand, one might be reluctant to admit such radically false appearances, perhaps because one felt that it would require unduly tricky and *ad hoc* manoeuvring. (Think about accommodating a use of 'Catherine' in (10) to talk about Catherine herself.) One might think that it must always be transparent when an expression has been mentioned (ambiguities apart). But one could still retain the form of the coincidence by developing a suitably technical and non-pragmatic sense of drawing attention to, or talking about, something – for example, referring to it, where this is a notion of reference determined by a theory of strict and literal truth conditions. (This would involve denying that somebody using (7) had, strictly speaking, said something true, though he or she may have succeeded in conveying a true thought.) It is no part of my contention to deny that one could take such tacks as these.

What one could not do is maintain both of the following: that the use/mention distinction is a clear-cut one; and that it has the significance usually attributed to it. If one took the first tack above and tied the distinction to what people are talking about in the relevantly vague and pragmatic sense, then the use/mention distinction would itself be correspondingly indeterminate. Consider: Jim remembers almost nothing of the physics he once knew, but he still knows that electrons are elementary particles with negative charge. Did I just make a claim about electrons? or about the word 'electrons'? or both? What was I drawing attention to when I employed the word 'electrons'? Is there a fact of the matter? If not, there is not fact of the matter, on this approach, whether I just mentioned the word 'electrons'.

If one took the second tack and introduced a technical sense of talking about things, then one would remove the use/mention distinction from the arena of pragmatics and also, thereby, from the arena where the confusions to which it was supposed to be pertinent arise. For example, as a matter of pragmatics, one could use (8) to convey a truth about certain arithmetical expressions – a truth which arguably could not be conveyed in any other way – even though, on this approach, one would not have talked about those expressions in the relevantly technical sense and might indeed have said something literally false. The fact remains that one would be suffering from the kind of confusion that is of concern here if one's purpose in thus using (8) was, say, to undermine the thesis that it is impossible to think something that is impossible, such as an arithmetical absurdity. (The relevant kind of confusion has to do with what one is talking about in a pragmatic, non-technical sense.) This approach therefore has the consequence that the use/mention distinction does not impinge directly on such confusion and so does not have that particular significance, precisely the significance which it is usually thought to have.

III

I have said that if there is a range of easily recognizable and specifiable devices which register when an expression has been mentioned, then the fact that an expression has been mentioned cannot be related in any neat way to what it has been employed for. Corroboration for this comes from further reflection on what one of the relevant devices would be: quotation marks. Davidson suggests that we read a pair of quotation marks as 'the expression a token of which is here' ([1], p. 90). An expression is supposed to be a purely syntactical entity. But surely there are times when it would be more appropriate to read a pair of quotation marks as 'the word a token of which is here', where a word is thought of as having a semantic component. Consider, for example, saying,

(13) 'Transcendental' is an esoteric word

or

(14) It would be an insult to call somebody 'clever'

not meaning thereby that it would be an insult to use that *expression.* Indeed sometimes quotation marks are used to pick out something *purely* semantic, as in a typical use of

(15) The answer to this question is 'No'

where one might just as well have used

(16) The answer to this question is no.

Davidson could obviously admit that quotation marks are capable of being used in these ways, by insisting that they are ambiguous; and he might go on to say that only by using quotation marks in the way he describes does one strictly mention an expression. But this would be sufficiently *ad hoc* to require justification. (It would not be outrageous, however. We might want to single out a distinct use of quotation marks for, say, referring to books by their titles.) A likelier proposal, which could still be offered within the framework of a demonstrative theory, is that quotation marks resemble such indexicals as 'this' and 'that' in leaving open a wide range of things that they can be used to demonstrate, the demonstratum on any occasion of use being fixed by appeal to context (*cf.* [3], p. 3 and pp. 16–17, and [4], p. 4). If this is roughly correct, the possibility arises that they can be used, in particular, to demonstrate actual employments of expressions, or meaningful utterances, and this in turn suggests that they have a use very much like the use of 'that' which Davidson outlines in his paratactic account of indirect discourse in [2]. On this construal, there could be a perfectly legitimate use of

(17) 'Grass is green' is true

which had exactly the same import as a typical use of

(18) It is true that grass is green.

And it seems to me that there could be. However that may be, we see here once again that what one is employing an expression for when one mentions it is heavily context-dependent and capable of merging imperceptibly into what one is employing it for when one uses, but does not mention, it – provided that it is appropriately transparent when an expression has been mentioned. (If it is not, use and mention themselves merge imperceptibly into each other.) This reinforces my main contention. But it also casts doubt on a rather glib appeal to the use/mention distinction which is frequently made in the philosophy of language, and I shall close by making some brief comments about that.

Consider the two following sentences:

(19) Mary knows that '$e^{i\pi} + 1 = 0$' is true
(20) Mary knows that $e^{i\pi} + 1 = 0$

It would be typical of what I have in mind to insist, simply on the strength of the use/mention distinction, that somebody who asserted both of these sentences would be making two very different claims, the difference being one which (for familiar reasons) has a vital bearing on questions of understanding. I do not want to deny that these sentences could be used to make two different claims. But I do want to say, in the light of the discussion above, that they could equally well be used to make the same claim. So provided we can be confident that somebody asserting (19) would be mentioning the mathematical expression, and that somebody asserting (20) would not be, then simply appealing to this fact does not ensure that there would be any difference between the two claims which the person was making. Consequently, if we do want to consider uses of (19) and (20) to make two different claims, much more needs to be said about what we are considering and what philosophical significance it is supposed to have than is provided by some cursory reference to the use/mention distinction. This is especially significant because of a certain tension which will be felt (a tension which is typical of what will be felt in other similar cases). The firmer a grip we have on the difference between the two claims, then the more we shall be viewing the linguistic entity to which Mary is being related by the first claim as an abstract and theoretical construct (a sequence of typographical shapes, say) and the less we shall think of it as something which ordinarily impinges on people who handle the language of mathematics with understanding, so that the less, in turn, we can be confident that what we are considering is able to throw much light on what it *is* to handle the language of mathematics with understanding. A similar tension will be felt 'from the bottom up' with respect to the second claim. For we shall feel some pressure to view it in such a way that it does not treat of Mary's linguistic practices or of what she knows about the language of mathematics, and this is, to put it mildly, not easy. Be that as it may, we are brought to realize that here too the use/mention distinction might not have quite the significance which it has customarily been supposed to have.[1]

King's College, Cambridge CB2 1ST © A. W. MOORE 1986

REFERENCES

[1] Donald Davidson, 'Quotation', reprinted in *Inquiries into Truth and Interpretation*, Oxford University Press, 1984.

[2] Donald Davidson, 'On Saying That', reprinted in *Inquiries into Truth and Interpretation*, Oxford University Press, 1984.

[3] L. Goddard and R. Routley, 'Use, Mention and Quotation', in *The Australasian Journal of Philosophy*, 44, 1966.

[4] Laurence Goldstein, 'Quotation of Types and Other Types of Quotation', in *Analysis*, 44, 1. January 1984.

[5] C. H. Whiteley, 'Names of Words: A Note on *Analysis* "Problem" No. 10', in *Analysis*, 17, 5. April 1957.

---

[1] My thanks are due to Anita Avramides, Naomi Eilan, Andrew Rein and an anonymous referee.

INDETERMINACY OF IDENTITY

By MONTE COOK

GARETH EVANS ('Can There be Vague Objects?', ANALYSIS, 38.4, October 1978, p. 208) and Nathan Salmon (*Reference and Essence*, Princeton 1981, pp. 240-52) have offered variations of an apparently simple and straightforward argument that seems to tell against the possibility that objects can be indeterminately identical. In what follows, I want to argue that this argument fails and that, despite the efforts of Sidney Shoemaker and H. W. Noonan, various purported examples of indeterminately identical objects have yet to be explained away. I want then to argue that the conclusion of the argument is nevertheless true and to explain away the seeming counterexamples.

I THE ARGUMENT

The basic argument is a *reductio*. Let '*a*' and '*b*' be singular terms and the sentence '$a = b$' be of indeterminate truth value – that is, neither definitely true nor definitely false. *b* must then have the property of its being indeterminate whether it is identical to *a*. *a*, however, does not have this property, since surely *a* is definitely identical with *a*. It follows that since *a* lacks a property that *b* has, *a* cannot be identical with *b*. But then if *a* cannot be identical with *b*, then it is not indeterminate whether *a* is identical with *b*.

Though one is likely to balk at the notion that there are such properties as being-indeterminate-whether-it-is-identical-to-*a*, this part of the argument is, I think, impeccable. It would indeed be questionable if one inferred that an object had such a property merely from the existence of an indeterminate identity statement. But the argument infers such properties only for cases in which the indeterminacy of the identity statement results from an indeterminacy in the identity of the objects referred to in the statement. The argument does not deny that identity statements can have indeterminate truth values, only that objects can be indeterminately identical. If the objects are indeterminately identical, then we should be able to conclude that each has the property of being indeterminately identical to the other. Of course, the Evans–Salmon argument is a *reductio*; and since Evans and Salmon believe that there is no such thing as indeterminate identity, they themselves presumably believe that there are no such properties. Nevertheless, anyone who maintains not merely that there are identity *statements* that are indeterminate but also that there are *objects* that are indeterminately identical seems committed to such properties.

II CRITICISM OF THE ARGUMENT

As John Broome points out ('Indefiniteness in Identity', ANALYSIS 44.1, January 1984, pp. 6–12), the argument errs in supposing that since *a* and *b* differ in their properties they must be non-identical. It's true that if *a* and *b* differ in their properties they must not be identical. With the initial assumption that *a* and *b* can be indeterminately identical, however, they can fail to be identical in two ways: either by being non-identical or by being indeterminately identical. If the property that distinguished *b* from *a* were an ordinary property – one not defined in terms of indeterminacy – we could conclude that *a* and *b* were non-identical. The property that distinguishes them, however, is not an ordinary one. On pain of begging the question, we cannot rule out the possibility that *a* fails to be identical with *b* because they are indeterminately identical. In fact, since the property that *b* has (and *a* lacks) entails that they are indeterminately identical, we should conclude that it is in *this* way that *a* fails to be identical with *b*.

Broome's criticism assumes that *a* and *b* can be in one of three mutually exclusive states: they can be identical, they can be non-identical, or they can be indeterminately identical. In contrast, in reasoning that since it is not the case that *a* is identical with *b*, *a* must be non-identical with *b*, the original argument assumes that there are only two such states: either *a* and *b* are identical or they are non-identical. Any defender of indeterminate identity worth his salt would naturally take cases of indeterminate identity as exactly the cases in which this assumption breaks down.

III INDETERMINACY OF DENOTATION

Broome's criticism does not establish that objects are sometimes indeterminately identical. We might try to establish this by examples like the following one from Sydney Shoemaker (*Personal Identity*, Blackwell 1984, 146):

> If ... the vagueness of the term 'building' makes it indeterminate whether Alpha Hall and Beta Hall should count as parts of the same building (they are connected by a rather flimsy covered walkway) it may be indeterminate whether the identity statement 'the building in which Smith is lecturing is the same as the building in which Jones is lecturing' is true, if Smith is lecturing in Alpha Hall and Jones is lecturing in Beta Hall.

Shoemaker denies that this sort of example shows that objects can be indeterminately identical, because he thinks that the real source of the indeterminacy here is that it is indeterminate which of several things 'the building in which Smith is lecturing' refers to. He fails to explain, however, what the several things in question are. Though it is indeterminate whether 'the building in which Smith is lecturing' refers to the building in which Jones is lecturing or one distinct from it, we cannot conclude that the reference of this phrase is indeterminate between a building in which Jones is lecturing and a building distinct from this. There aren't two buildings for 'the building in which Smith is lecturing' to refer to; there is only the one building in which he is lecturing, which is neither definitely identical nor definitely non-identical to the building in which Jones is lecturing.

Shoemaker seeks to explain all indeterminate identity statements on the model of identity statements with definite descriptions that lack a definite denotation. To make this explanation good, he has to show that such an ambiguity can occur both with names and with definite descriptions that by ordinary standards have a definite denotation; and to do this he must find competing entities for these names and definite descriptions to refer to. As it stands, Shoemaker's explanation is incomplete.

Adopting the same model, H. W. Noonan ('Vague Objects', ANALYSIS 42.1, January 1982, pp. 3-6) solves the problem by taking the names and definite descriptions in question to be indeterminate in their denotation between different 'summations' of spatio-temporal parts. Thus, 'the building in which Smith is lecturing' could denote a collection of parts encompassing just the Alpha Hall side of the walkway or it could denote a collection of parts encompassing both sides of the walkway; in fact, it fails to denote either one determinately. Similar comments apply to 'the building in which Jones is lecturing,' and the lack of a determinate denotation for the definite descriptions involved explains the indeterminacy of the statement that the building in which Smith is lecturing is identical to that in which Jones is lecturing.

The attractiveness (for some) of Noonan's approach lies in the thought that (a) there are different collections of spatio-temporal parts available and (b) buildings just are appropriately unified collections of spatio-temporal parts. Since the parts of a collection must be appropriately related for it to constitute a building, we can have different candidates for the building in which Smith is lecturing without committing ourselves to different *buildings*. The collections of parts are united in different ways; and for some collections, the concept of 'building' is not sufficiently precise to determine whether the collection is united by the appropriate relation.

As Noonan notes, those philosophers who find unappealing the notion of spatio-temporal parts or the notion that ordinary objects are summations of such parts will object to his explanation of indeterminate identity statements. Moreover, something seems wrong when a problem of understanding identity statements pushes us to metaphysics.

My basic criticism, however, is not that Noonan's metaphysics is faulty, but that the basic model that both he and Shoemaker use to explain apparent cases of indeterminately identical objects just seems to involve a confusion. It is indeterminate whether 'the building in which Smith is lecturing' denotes a building encompassing just Alpha Hall or one encompassing both Alpha and Beta Halls; by an appropriate sharpening of the notion of a building we could make this determinate one way or the other. Saying this, however, misleads in that it tempts us to locate the indeterminacy in the wrong place. The denotation is not indeterminate; the boundaries of what is denoted are indeterminate. Obviously enough, 'the building in which Smith is lecturing' denotes the building in which Smith is lecturing. Though it is indeterminate whether it denotes a building encompassing just Alpha Hall or one encompassing both Halls, this phrase does not indeterminately denote 'the' building (or possible building or object that is possibly a building) that contains Beta Hall and 'the' building (or possible building, etc.) lacking it. There aren't two candidates for the building in which Smith is lecturing; there are two possible characterizations of the one building. When we sharpen the notion of a building, we are not making the denotation of 'the building in which Smith is lecturing' determinate; we are making it determinate whether this building includes Beta Hall. (If the colour of my car is on the borderline between red and orange, then it is indeterminate whether the colour of my car is red. 'The colour of my car,' however, has a determinate denotation. If we sharpen the notion of being red, we don't make the denotation of 'the colour of my car' determinate, we make it determinate whether the colour of my car is red.)

IV INDETERMINATE IDENTITY AND THE LAW OF EXCLUDED MIDDLE

So far I have been criticizing arguments against the position that objects can be indeterminately identical: the Evans-Salmon argument fails, and the attempt to use indeterminacy of denotation to explain away apparent cases of indeterminately identical objects is misguided. Nevertheless, I want to deny that objects can be indeterminately identical.

The key to arguing against indeterminately identical objects is to see how one can and should defend the Law of Excluded Middle and explain away apparent cases of indeterminate identity. The Law of Excluded Middle (LEM) says that every disjunction of the form '*P* or not-*P*' is true. We have seen how the Evans-Salmon argument goes awry in assuming without argument an instantiation of this. One way of stating the issue of whether indeterminacy is in the world is in terms of LEM: if indeterminacy is in the world then LEM does not hold – objects can be in three mutually exclusive states: they can be identical, they can be non-identical, and they can be indeterminately identical. To defend LEM, then, is to attack the view that objects can be indeterminately identical.

The first step in defending LEM from apparent counterexamples involving indeterminately identical objects is to distinguish LEM from the Principle of Bivalence (PBV), which says that every statement is either true or false, and to argue that the counterexamples tell against PBV rather than against LEM. The alleged cases of indeterminate identity and more generally various cases involving vagueness do indeed force use to give up PBV: it is neither true nor false that the building in which Smith is lecturing is the building in which Jones is lecturing, and neither true nor false that a car is red when its colour is on the borderline between red and orange.

Such cases, however, need not force us to give up LEM. Consider first the car whose colour borders between red and orange, so that the car is not definitely either one. It is natural to say that the car is neither red nor not red, and thus to think that the case forces us to give up LEM. If we keep clear on the distinction between LEM and PBV, however, we can see that though it is neither true that the car is red nor true that it is not red, it is nevertheless true that it is red or not red. To see this, it may help to think of 'red or orange' as carving out a colour space. It is true that the car is red or orange, even though no further specification can be given. That is, the car is red or orange – it is in this colour space – and not (say) green or blue, even though, because of the vagueness of the notions of *red* and *orange*, we cannot say either that the car is red or that it is orange. If, however, the car is either red or orange, it is either red or not red; and the case does not force us to give up LEM (cf. Michael Dummett, 'Wang's Paradox,' *Synthese* 30 (1975), pp. 301-24).

The trick to seeing that borderline cases do not force us to give up LEM is (a) seeing that a disjunction can be true even though neither disjunct is true, and (b) seeing that this does not prevent the disjunction's being exhaustive. The car is red or not red, even though it is neither true that it is red nor true that it is not red. Nevertheless, although the car satisfies a third alternative to being determinately red or determinately not red, it does not satisfy some third alternative to being red or not red.

We may be reluctant to say that the car is either red or orange (and thus either red or not red) for various reasons. For one thing, saying this may suggest that one of two clear-cut options is met, that the car is either definitely red or definitely orange. Here we need to distinguish its being determinately red or orange from its being indeterminate whether it is red or orange. Further, some inferences like disjunctive syllogism that usually work will fail: though it is true that the car is either red or orange and not true that the car is red, it does not follow that the car is orange. Here I think we have either to qualify disjunctive syllogism or to say that it sometimes just fails. Finally, one will not be able to explain the truth conditions for the statement that the car is red or orange in a straightforward truth-functional way. Here I think one should appeal to something like true-under-an-acceptable-sharpening. Thus, one might say that 'red or orange' is true of an object if under all acceptable sharpenings of 'red' and 'orange' either 'red' or 'orange' is true of it (cf. Dummett, *ibid*).

Borderline cases of colour show us how '*P* or not *P*' can be true even though neither disjunct is true. Now consider Shoemaker's case and begin by considering where Jones is. Though he could be in various other places, in fact he is either in the building in which Smith is lecturing or in a separate building containing Beta Hall. The vagueness of what counts as being part of the same building prevents our further specifying which of these is the case. Nevertheless, if he is either in the building in which Smith is lecturing or in a separate building, then he is either in the building in which Smith is lecturing or he is not. If he is, then the building in which Smith is lecturing is identical to that in which Jones is lecturing; if he is not, then the building in which Smith is lecturing is not identical to that in which Jones is lecturing. In sum, though the vagueness of what counts as being part of the same building prevents our determining whether the building in which Smith is lecturing is identical to that in which Jones is lecturing, either it is or it isn't, and we are not forced to give up LEM.

The source of the indeterminacy of 'the car is red' is the vagueness of the notion of *red* – the absence of a sharp line between being red and not being red. The source of the indeterminacy of 'the building in which Smith is lecturing is identical to that in which Jones is lecturing' is the vagueness of *part of the same building*, but this case is slightly more complicated in that the role of the notion of *part of the same building* is concealed. The indeterminate statement does not contain the expression 'part of the same building.' Indeed, it need not even contain the expression 'building': we can name the buildings '*a*' and '*b*' and speak of the indeterminacy of '$a = b$.' This last fact tempts one to think that the indeterminacy must somehow result from an ambiguity in '*a*' or '*b*' (or perhaps in '=') – the vagueness of 'part of the same building,' it can seem, must somehow be captured by an ambiguity in one of these.

In fact, however, the vagueness of this notion can infect the statement that '$a = b$' without any ambiguity in the component symbols of this statement. The problem is not that one (or more) of these symbols is ambiguous, but that the vagueness of 'part of the same building' leaves it open whether $a = b$ or not. To see how this works, consider an analogy. Suppose that two cars, *A* and *B*, have colours that are on the borderline between red and orange and that with one way of drawing a precise line between red and orange the cars have the same colour, while with another way of drawing it they don't. It is indeterminate, then, whether the two cars have the same colour. Nevertheless, (1) either they have the same colour or they don't – though neither one of these is definitely the case, there is no third possibility. Further, (2) 'the colour of *A* (*B*)' has a determinate denotation, viz., the colour that *A* (*B*) has. Moreover, (3) 'same colour' is not ambiguous – or at least, any ambiguity in 'same colour' is not the source of the indeterminacy. Even with no ambiguity at all in what counts as having the same colour, it is indeterminate whether *A* and *B* have the same colour. The problem of course, is that although *A* and *B* would have the same colour if both were red, the vagueness of 'red' makes it indeterminate whether both are red, and hence indeterminate whether they have the same colour.

Shoemaker and Noonan appealed to indeterminacy of denotation as a way to save LEM in light of vague identity statements. By the Shoemaker-Noonan position, if '*a*' and '*b*' determinately denote objects, then

(1) it is true that either $a = b$ or $a \neq b$; and
(2) it is either true that $a = b$ or true that $a \neq b$.

If this appeal had worked, we could have said that the objects weren't indeterminately identical, the statements were indeterminate due to an indeterminacy of denotation. In the present section I have tried to save LEM by distinguishing between LEM and PBV and arguing that (1) can be true even though (2) is false. The explanation of vague identity statements lies neither in the LEM-violating indeterminacy of the objects referred to in these statements nor in the indeterminacy of the denotation of certain components of these statements, but in our failure to draw a sharp line between what counts as part of the same object and what doesn't.[1]

University of Oklahoma,
455 West Lindsey Street, Norman,
Oklahoma 73019, U.S.A.

© MONTE COOK 1986

---

[1] I wish to thank Chris Swoyer for many valuable discussions on the topic of vague identity.

ON SETS AND WORLDS: A REPLY TO MENZEL

By PATRICK GRIM

IN [4] I offered a Cantorian argument that there can be no set of all truths, and noted one application: against an approach to possible worlds in terms of maximal consistent sets of propositions. Shortly thereafter Selmer Bringsjord offered a similar argument against set-theoretical worlds in [1].

In [7], a recent and important contribution to the discussion, Christopher Menzel has raised a number of critical points regarding Bringsjord's argument. But these points also apply against my earlier and more general argument. I want to address them directly here.

I

There cannot, it appears, be a set of all truths. For essentially the same Cantorian reasons, it seems, there can be neither world-stories nor worlds construed as maximal consistent sets of propositions.

But perhaps there *can* be a non-set *something-else* of all truths, or a non-set something-else appropriate to the propositions of a world-story or a world.

Menzel does ultimately have such a something-else to offer. He first considers and rejects, however, what may seem the obvious candidate: the proper classes of Von Neumann-Bernays-Gödel (VNBG) and similar systems.

In [7] as in [8] Menzel deftly distinguishes two conditions on proper classes: excessive size and unbounded rank.[1] It is the second he takes as the 'true conceptual boundary' between sets and proper classes. If propositions are admitted as urelements, however, a world-story S will be of *bounded* rank:

> since ... propositions for the world-story theorist are not sets, they themselves have no members and so have a rank of 0; hence, S itself has a rank of 1, i.e. S is bounded. ([7], p. 70)

There is then no need to consign world-stories (or a totality of truths) to the realm of proper classes. Or so the argument goes.

Note however that Menzel's 'ranks' here are confined exclusively to sets: it is sets alone which have ranks, determined exclusively by the ranks of their members. But this may be a dangerously provincial view of rank. Hierarchies of types have of course often been thought necessary for propositions, predicates, and properties as well as sets, and in the end an *integrated* theory of types may be required. We may find it necessary to rank sets of propositions above those sets mentioned in propositions they include, for example.[2] But given an integrated ranking of this sort Menzel's argument fails to go through. It is no longer clear, in particular, that world-stories will have a rank either of 1 or of any other bounded *n*.

Nonetheless there *are* strong reasons to think that Menzel's *conclusion* here is right: that proper classes won't do as a something-else appropriate to world-stories or a totality of truths.

---

[1] Within VNBG proper classes are those too large to be sets. Menzel defines the *rank* of an object recursively as the least ordinal greater than the ranks of all its members. A class is *unbounded* if it contains members of arbitrarily high rank.

[2] This would seem the obvious way to deal with the following variant of Russell's paradox, for example:

Some sets of propositions contain propositions which mention those sets themselves. Some sets of propositions contain no such propositions. We will refer to the latter as 'non-self-mentioning' propositional sets.

Consider now the set *S* of all propositions which mention only non-self-mentioning sets. Is *S* self-mentioning or not?

Some of the objections to proper classes are technical. Within VNBG and similar systems, for example, proper classes are kept immune from Cantorian arguments only at the cost of sacrificing general provability for instances of the induction schema (see for example [6] p. 198).

Intuitive objections are even stronger, however. For in VNBG and similar systems proper classes are introduced precisely as classes which are not members of further (even finite) classes. That would clearly not seem to hold for either a 'class of all truths' or for classes appropriate to world-stories. For surely a class of all truths *would* be a member of further classes – of various classes of propositional classes and various classes of classes of truths. Surely a class of all truths would form a pair with a class of all falsehoods, for example. Worlds and their world-stories would seem if anything *less* amenable to treatment as proper classes. For consider the class of worlds in which I exist and the class of worlds in which this vase is hit with a hammer. Current philosophical use quite generally seems to demand *classes* of possible worlds.

II

The 'something-else' that Menzel ultimately suggests involves a recourse not to amplified but to amputated set theories. Perhaps the world-story theorist, though clearly a profligate realist regarding propositions, can follow an abstemiously constructivist line when it comes to sets:

> By adopting ZF-Power (or some similarly conservative set theory), then, and adjusting it appropriately to allow for the existence of large sets (and urelements), the world-story theorist is free to postulate the existence of his world-stories without fear of paradox, at least by way of [the Cantorian] argument; for that argument depends essentially on there being a power set of the world-story S; but there is simply no way of generating the full classical power set of an infinite set in ZF-Power... ([7], p. 71)

It is far from clear, however, that large sets within ZF-Power will prove any more satisfactory here than proper classes.

Not too surprisingly, sacrifice of the power set axiom results in quite major technical limitations; at one blow we are effectively exiled from Cantor's paradise (some surprising technical limitations of ZF-Power appear in [10]). Nor does ZF-Power applied as Menzel envisages it really seem to satisfy constructivist scruples. Menzel relies on the fact that 'there is simply no way of generating the full classical power set of an infinite set in ZF-power...' But of course there is no way of *generating* basic infinite sets within ZF-Power at all; they must be added by fiat of special axiom. The set of zero and its successors appears by axiom within ZF, and a set of all ordinals is added by special axiom in Menzel's system in [8]. If we are to include the reals, and world-stories, and infinite sets of world-stories within ZF-Power we will need special axiomatic provision for these as well – a prodigal postulation of sets far from constructivist in spirit.

It is not clear, moreover, that sacrifice of the power set axiom alone will be enough to escape Cantorian difficulties regarding a set of all truths. For consider the following argument.

Let us suppose we *did* have a set T of all truths within ZF-Power. Then by the axiom schema of separation, carried over directly from ZF, we would have as a subset of T a set of all truths satisfying a particular condition $B(x)$. As long as our basic language is rich enough to express the notion that a truth t is about a topic c, it appears, one such condition will be 'x is about a set of truths.' By separation, then, we would have a set C of all truths about sets of truths.

More formally, using '$\mathrm{A}xy$' to indicate that x is about y, our condition 'x is about a set of truths' is:

$$\exists y \forall z (z \in y \supset z \in T \cdot \& \ \mathrm{A}xy).$$

The axiom schema of separation, taken directly from ZF, is

$$\forall z_1 \ldots \forall z_n \forall a \exists y \forall x (x \in y \equiv x \in a \ \& \ B(x)),$$

where $z_1, \ldots, z_n$ are the free variables of $B(x)$ other than x, and the only restriction on $B(x)$ is that it does not contain y as a free variable (see for example [6] p. 175). Using T for a in this schema and our condition above for $B(x)$ would give us a set C of all truths about sets of truths:

$$\forall x (x \in C \equiv x \in T \ \& \ \exists y \forall z (z \in y \supset z \in T \cdot \& \ \mathrm{A}xy)).$$

The existence of a set C, however, would give us a reductio. C would, in particular, be larger than T. For consider any one-to-one function f from T into C, mapping truths onto truths concerning sets of truths, and consider further a set c' of all truths which do not belong to the sets their assigned element is about. Here using '$a(f(x))$' to indicate the set (or union of sets) $f(x)$ is about,

$$x \in c' \equiv x \in T \ \& \ x \notin a(f(x)).$$

Clearly C will contain some truth about c' just as it contains some truth regarding any set of truths. But by familiar reasoning f can assign no element of T to any truth about c'. C is larger than T.

One branch of the reductio, then, is this: given a set T of all truths, it appears, there would be a subset C larger than the set T of which it is a subset.[3]

---

[3] Note that this is not a problem for sets in ZF-Power in general; the argument above relies essentially on the supposition that T is a set of all *truths*.

Here and elsewhere I am obliged to an anonymous referee for helpful comments.

Another branch is this. Each element of C is a truth. There are thus more truths than elements of T, and thus T cannot, as assumed, be a set of *all* truths. A similar argument would of course apply to world-stories.

In order to escape the Cantorian difficulties of C, it appears, something more would have to be sacrificed – something above and beyond the power set axiom. Perhaps additional restrictions on the axiom schema of separation are called for. Perhaps the language of the system must exclude or restrict the notion of truths about sets of truths – though since there clearly *are* truths about any sets of truths this would seem far from satisfactory. Or perhaps here again we should simply conclude that there is no set T of all truths.

III

Variations on the Cantorian argument can be used to show that there is no set of all necessary truths, no set of all falsehoods, no set of all propositions, and no set of all things known by an omniscient being (see also [3] and [5]). If extensions are construed standardly in terms of sets, 'true', 'false', 'necessary', and 'is a proposition' have no extensions. Since the conjunction of a necessary and a contingent proposition is itself contingent, moreover, the argument can also be adapted to show there is no set of all *contingent* truths or falsehoods.[4] Nor is there any need to interpret the argument as restricted to linguistic entities in any sense; the argument would be the same against a set of all facts or all states of affairs.

In a concluding caveat, however, Menzel appeals to two outlines of possible worlds which appear to escape Cantorian arguments. One of these is Plantinga's:

> a state of affairs S is *complete* or *maximal* if for every state of affairs S′, S includes S′ or precludes S′. And a possible world is simply a possible state of affairs that is maximal. ([9], p. 45)

The other is Menzel's own:

> A world-story [could] be a set S that is maximal in the sense that for any proposition q, S *entails* (but not necessarily *contains*) either q or its negation, but not both ... ([7], p. 72)

---

[4] In the standard argument (as in [4]) we envisage a necessary truth corresponding to each element of the power set. Here we envisage the conjunction of each of these with a chosen contingent proposition.

It appears that even a set of all atomic propositions will fall victim. For suppose a set A of all atomic propositions and a set C of all connectives. Within classical atomism compounds from these will give us all propositions. If A and C are sets, however, by standard principles of set theory all permutations of elements of A sprinkled with permutations of elements of C will *also* form a set. We know independently that there can be no set of all propositions, however, and thus by modus tollens it appears there can be no set of all classical atomic propositions.

Do these approaches offer an escape from Cantorian difficulties of sets and worlds?

That is unclear. Menzel's outline demands universal quantification over propositions, Plantinga's calls for quantification over all states of affairs, and it appears that any genuine explication of the notion of 'maximality' crucial to possible worlds must be similarly quantificational in form. The lingering difficulty here is the following. By our Cantorian arguments there can be neither a set of all propositions nor a set of all states of affairs. But the only real semantics for *quantification* we have is in terms of sets.[5]

State University of New York,
Stony Brook, NY 11794, U.S.A.

© PATRICK GRIM 1986

REFERENCES

[1] Selmer Bringsjord, 'Are There Set-Theoretic Possible Worlds?', *Analysis* 45.1, January 1985, p. 64.
[2] Martin Davies, *Meaning, Quantification, Necessity*, London: Routledge and Kegan Paul, 1981.
[3] Patrick Grim, 'Some Neglected Problems of Omniscience', *American Philosophical Quarterly*, 20 (1983), 265-76.
[4] Patrick Grim, 'There Is No Set Of All Truths', *Analysis* 44.4, October 1984, pp. 206-8.
[5] Patrick Grim, 'Against Omniscience: The Case from Essential Indexicals', *Noûs*, 19 (1985), pp. 151-80.
[6] Azriel Levy, 'The Role of Classes in Set Theory', in Gert H. Müller, ed., *Sets and Classes*, Amsterdam: North-Holland, 1976, pp. 173-215.
[7] Christopher Menzel, 'On Set Theoretic Possible Worlds', *Analysis*, 46.2, March 1986, pp. 68-72.
[8] Christopher Menzel, 'On The Iterative Explanation of the Paradoxes', *Philosophical Studies*, 49 (1986), pp. 37.
[9] Alvin Plantinga, *The Nature of Necessity*, Oxford: Clarendon Press, 1974.
[10] Andrzej Zarach, 'Unions of ZF^--models which are themselves ZF^--models', in D. Van Dalen, D. Lascar, and T. J. Smiley, eds., *Logic Colloquium '80*, Amsterdam: North-Holland, 1982.

---

[5] In [2], p. 262, Martin Davies suggests further difficulties: a Cantorian argument against a set of all worlds, which would seem to apply however worlds are conceived. Crucial premises of Davies' argument, however, involve principles both messy and intriguing regarding philosophy of mind and the conceivability of propositions.

FALSE STIPULATION AND SEMANTICAL PARADOX

By LAURENCE GOLDSTEIN

WE cannot stipulate

(1) Strawson = Quine

but we can stipulate

(2) Seth = Pringle-Pattison

The reason is obvious: 'Strawson' and 'Quine' are names already assigned to quite different persons; 'Pringle-Pattison' was unassigned until, in order to satisfy the terms of a will, it was assigned to the unfortunate Andrew Seth. (2) is true, (1) is false.

We can stipulate

(3) $x = y + 1$

so long as 'x' and 'y' are not both previously assigned but we cannot stipulate

(4) $z = z + 1$

in standard arithmetic, since no number can be equal to its successor.

In [5], I gave a series of simple arguments showing that one cannot stipulate

(5) k = 'k is not true'

Because (5) is used to generate one version of the Liar paradox, that paradox can be dissolved, so I wished to show, by seeing that the contradiction to which the Liar paradox leads stems from a false stipulation. If, in doing some algebra, one arrived to one's consternation at a contradiction, then if, in tracing back through the reasoning, one discovered that one had inadvertently assumed (4), that would be quite sufficient to dissipate the puzzlement: it is no surprise that a contradiction follows from a necessary falsehood.

In [2] Buckner and Smith agree that, if the symbols flanking the identity sign in (5) are read as specifying statement contents, then (5) is indeed false in virtue of the meaning of the truth predicate. They comment: 'Perhaps this is Goldstein's intended point'. Well, *of course* that was my intended point, and since, as I subsequently discovered, that point can readily be extended so as to solve both the Epimenides and the Curry Paradoxes (see [7]) I am encouraged to think that [5] constitutes a straightforward, wholesome and satisfying solution of the Liar. Buckner and Smith however think that 'the Liar remains very much in business', on account of two objections they raise.

I argued in [5] that one can be misled into believing that (5) is a legitimate stipulation because, on the classical (Tarskian) theory of quotation, the 'k' on the right hand side of (5) is just part of a semantically unstructured name – it has no independent semantic significance, so that we are free to use 'k', on the left hand side as a name. The Tarskian theory is, however, mistaken (see [3], [6]) and once this is appreciated, the way is clear to recognizing the falsity of (5). Buckner and Smith contend that, even if the quotation marks in (5) are not Tarskian, we can *introduce* a quotational device that does function in a Tarskian or neo-Tarskian way. Let this quotational device by symbolized by a pair of asterisks. Their proposal is that '*k is not true*' is to function as a structural description of the sequence of symbols: k is not true. They then claim that

(6) *k* = **k* is not true*

legitimately stipulates that 'k' also denotes that sequence, and that the Liar paradox has thereby been reinstated.

Now, (6) is an entirely legitimate stipulation, but it is a stipulation about a sequence of symbols. Sequences of symbols do not have truth values, hence no semantic paradox can be derived from (6). The result of concatenating 'I' ' ' 'a' 'm' ' ' 'y' 'e' 'l' 'l' 'o' 'w' is just a series of marks – a grapheme (see J. L. Austin's distinctions in Lecture VIII of [1]). It is not a truth about a bowl of custard nor a falsehood about a Chinese person nor the confession of a coward. Similarly, the *type sentence* 'I am yellow' is neither true nor false. Through a misreading of my paper 'Quotation of Types and Other Types of Quotation' [4] Buckner and Smith assume that I would read (5) as

(7) k = the type of which a token is displayed here: k is not true.

But the quotation in (5) is not the quotation of a type but, as mentioned earlier, the quotation of a statement-content. Only when what is quoted in (5) is something that has a truth value do we embark on the Liar reasoning. But when this *is* how the quotation marks in (5) are read, then we can see that (5) is false so that the paradoxicality evaporates.

Buckner and Smith's second objection is that bringing to bear a theory of quotation on the Liar is only effective (if it's effective at all) against those particular versions of the paradox which, like (5), employ quotation marks. There are, as they point out, many devices other than quotation marks by means of which the necessary self-reference can be achieved.

This is so, and if my solution applied only to direct-quotational versions of the Liar it would not be worth the paper it's written on. But in fact my solution is entirely general. The paradoxicality of the Liar arises from failure to detect a false stipulation. I chose to discuss a direct-quotational version of the paradox simply because, in that version, the stipulation and its falsity are relatively easy to uncover. My discussion of quotation was not intended as a solution to the Liar but as support for a hypothesis about why the presence of these elements in the quotational version had not been detected sooner.

I find it a little puzzling that this second objection should even have been raised. Shortly after writing up my solution in [5] I came across G. H. von Wright's *Philosophical Logic* [8] where a solution to the Liar along lines very similar to mine is proposed. But von Wright discusses not a quotational version of the paradox but the more common formulation: 'This is not true'. And Smith published a review of von Wright's book, so he, at least, should have been aware that the second objection is misguided.

If one statement says of another that it is not true, then the statement referred to must have the opposite truth value to that of the statement referring to it, and so must be a *different* statement. Hence a speaker can no more stipulate that these statements are identical than he can stipulate that a number be identical to its successor. But in *any* self-referential version of the Liar the assumption is made that there *can* be *one* statement which refers to that same statement and says of it that it is not true. In most versions of the paradox, this assumption does not take the form of an explicit stipulation; it needs to be uncovered. Again, a numerical example may offer a useful comparison: Consider four numbers, a, b, c and d, where a, b and c are designated by the familiar numerals '2', '3' and '4' respectively. The number $d = 5$ if $\Sigma \neq 14$, where Σ is the sum of all four numbers; if $\Sigma = 14$, then $d = 6$. At first glance it looks as if d must be 5 or 6, since Σ must either be 14 or some number other than 14. However, a little calculation shows that *there can be no number d* satisfying the above conditions. In setting up the example, we referred to 'the number' d, and assumed that there must be such a number. That assumption, like the assumption underlying the Liar reasoning, is just mistaken.

Hong Kong University,
Pokfulam Road, Hong Kong

© LAURENCE GOLDSTEIN 1986

REFERENCES

[1] J. L. Austin, *How to do Things with Words*, Oxford: Clarendon Press, 1962.

[2] D. Buckner and P. Smith, 'Quotation and the Liar Paradox', *Analysis* 46.2, March 1986, pp. 65-8

[3] D. Davidson, 'Quotation', *Theory and Decision*, 11 (1979), pp. 27-40, reprinted in his *Inquiries into Truth and Interpretation*, Oxford: Clarendon Press, 1984.

[4] L. Goldstein, 'Quotation of Types and Other Types of Quotation', *Analysis* 44.1, January 1984, pp. 1-6.

[5] L. Goldstein, 'The Paradox of the Liar – A Case of Mistaken Identity', *Analysis* 45.1, January 1985, pp. 9-13.

[6] L. Goldstein, 'The Title of this Paper is "Quotation" ', *Analysis* 45.3, June 1985, pp. 137-40.
[7] L. Goldstein, 'Epimenides and Curry', *Analysis* 46.3, June 1986, pp. 117-21.
[8] G. H. von Wright, *Philosophical Logic*, Oxford: Basil Blackwell, 1983.

ON A SUPPOSED TEMPORAL/MODAL PARALLEL

By E. J. LOWE

IN 'On a temporal slippery slope paradox' (ANALYSIS 46.1, January 1986), D. E. Over attempts to formulate a temporal paradox of artefact identity paralleling Chisholm's well-known modal paradox. I do not believe that the parallel works but think that it fails interestingly because of an important dissimilarity between temporal relations and relations between possible worlds, consisting in the fact that the former put instants of time into a unique linear order not paralleled by any analogous linear ordering of possible worlds. In short, time constitutes a *dimension* in a way that possibility does not.

Chisholm's paradox may be posed in the following manner. Suppose we allow (as seems plausible) that an artefact could have originated from a set of component parts slightly different from those from which it actually originated but not from a wholly different set. Then we may be invited to consider a sequence of worlds in which a succession of slight changes in the originating components is envisaged in a fashion consistent with the identity of the resulting artefacts in successive worlds in the sequence, yet such that we end up with a world in which the originating components are completely different from those of the actual world, so that, by virtue of the transitivity of identity, we are seemingly obliged after all to identify the artefact resulting in this final world with the actual artefact – contrary to our initial supposition.

Over's supposedly parallel *temporal* paradox is alleged to arise from the following principle, which he takes to be analogous to the modal supposition just discussed:

> (T) If a given artefact x is assembled from a set of components y, then x could be completely disassembled and later re-assembled from another set of components z which sufficiently overlaps y; but x could not later be re-assembled from any set of components z' which does not sufficiently overlap y. (pp. 15–16)

In order to encourage us to see as plausible the second part of (T), Over invites us to imagine the following case:

> Suppose that we take our watch x to a watch-maker who completely disassembles it. He then gets its components muddled up with the components of many other watches, and at a later time t' he puts a watch x' together which does not contain many of the components of our watch. We naturally reject x' as not our watch. (p. 16)

And in order now to generate the supposed paradox Over describes a second case as follows:

> Perhaps our watch [has] to go back to the watch-maker many times for disassembly and reassembly with a small change in component parts. In the end we might reach a time t_n when we get a watch x_n made of exactly the same components as x' put together in the same way, but which should be our watch x if the first part of (T) is applied throughout this process. (p. 16)

But really all that Over has done here is to make it clear why the intuitions which sustain our denial of an identity between x and x' in his first case are not in fact adequately articulated by his principle (T). What he has shown by inviting us to consider the possibility of there being a temporal *sequence* of disassembly/reassembly episodes culminating in the assembly of a watch x_n which we identify with our watch x despite the considerable change in components is quite simply that the second part of (T) is *far too strong*, and is *not* sustained by the intuitions which lead us to deny the identity of x with x' in Over's first case. What his second case shows is precisely that, contrary to the second part of (T), *x could* later be reassembled from a set of component parts z' having *very little* overlap with the original set of component parts *y provided that* this state of affairs were to be arrived at through just such a sequence of episodes. And what his *first* case shows is that if this same degree of change in components were to take place *in one fell swoop*, rather than through such a sequence, then indeed the identity of our watch x would *not* be preserved.

In short, Over has not presented us with anything really *paradoxical* at all, since in a genuine paradox all the assumptions from which a conflict arises should have a strong claim on our acceptance – which (T) plainly does not. Yet the modal analogue of (T) *does*, I suggest, have quite a strong claim on our acceptance, which is why Chisholm's 'paradox' is worthy of the name. Why, then, does the analogy break down? Just for the following reason, I should say. We could see what was wrong with (T) by observing that it failed to distinguish, in its second part, between effecting a large change in components *in one fell swoop* and effecting it *through a sequence of minor changes.* Thus (T) could and should be amended precisely by insisting that any large change in components is only consistent with an artefact's continued identity provided that it is effected through a temporal sequence of minor changes (a point also insisted on by Hume in the *Treatise* I iv 6). But how could we correspondingly 'amend' the modal analogue of (T)? The fact is that we *could not*, in any principled way, because possible worlds are not linearly ordered along a dimension of 'possibility' in anything like the manner in which instants are linearly ordered in time. It is true enough that we talk of some possibilities being more 'remote' than others – but this is a very loose and imperfect metaphor at best. In Chisholm's paradox the 'sequence of worlds' we are invited to consider only constitutes a linear ordering by virtue of the stipulation that the position of each world in the sequence is determined by the degree to which the set of components in that world differs from the set existing in the actual world. There is, apparently, no *modal* comparative relation between these worlds which *independently* puts them into any linear order whatever. (Relative modal 'accessibility' seems to me at best merely a logician's fiction.) But in the seemingly parallel temporal case we can put into a linear order the *times* at which the different sets of components are assembled not *merely* by reference to the degree to which these sets differ in their membership from the original set, but *also*, quite independently, by reference to the *temporal* relations between these times (i.e. the degree to which one time is *earlier* or *later* than another). The required amendment to (T) will have the effect of allowing that artefact identity may be preserved across time at least where these linear orderings *coincide*, i.e. where the degree of difference in components gradually increases with the elapse of time. In the modal case however no analogous coincidence can be spoken of, nor hence allowed for by way of amendment.

University of Durham, 50 Old Elvet,
Durham DH1 3HN.

©E. J. LOWE 1986

IS THERE A TEMPORAL SLIPPERY SLOPE PARADOX?

By D. E. OVER

IN my paper, 'On a Temporal Slippery Slope Paradox' (ANALYSIS 46.1, January 1986), I argue that there is a temporal paradox which parallels a much discussed modal slippery slope paradox, often called Chisholm's paradox. But E. J. Lowe in his paper, 'On a Supposed Modal/Temporal Parallel' (ANALYSIS 46.4, October 1986), denies that there is a temporal paradox like the modal one. His argument is that principle (T), which I use to derive the supposed temporal paradox, does not really 'have a strong claim on our acceptance', while the corresponding modal principle does.

Lowe draws attention to the second part of principle (T). This states that an artefact x cannot be disassembled into its set of components y and then later be re-assembled from a set of components z which does not sufficiently overlap y. Lowe points out that this part of (T), as it stands, appears to be too strong in light of an example I use. In this example, the artefact x goes through a series of stages in which it is disassembled and re-assembled many times, but with a small change in its component parts each time. Could the series end with the re-assembly of x using components z? Perhaps it could. Suppose that the series takes place over a long period of time as the basic components of x slowly wear out, and at each stage a small number of worn out components are replaced by new ones. This case would then become similar to that of the ship of Theseus, except for the complete disassembly and re-assembly at each stage, and there would certainly be some reason for saying that it showed the second part of (T) to be too strong.

On the other hand, consider an example similar to the one I use first in my paper. We take the watch x to a watch-maker who disassembles it into its component parts y. He takes a very small number of these parts and adds them to many other parts he already has to form a new collection of parts z which overlaps very little with y, and then he assembles a watch x' from z. Could this watch x' be our watch x? Our intuition is strong, I think, that it could not be, especially if the many components displaced in this process were perfectly serviceable and perhaps brand new. Would we accept that x' was our watch if the watch-maker told us that he had quickly performed a series of assembly and disassembly operations, moving by such stages from x and its components y to x' and its components z? No, I am sure that we would call his performance a charade and still reject x' as our watch x. Sometimes we feel that the identity of an artefact is preserved under a series of assembly and disassembly operations, with a small change in components at each stage, and sometimes we do not.

I did express doubt about the exact formulation of (T) in my paper, but also noted how strongly we feel that some (T)-like principle is true (pp. 16–17). The problem is that any attempt to state a true (T)-like principle is threatened with inconsistency. Suppose that we weaken the second part of (T) by saying that x cannot be re-assembled from a set of components z which does not sufficiently overlap y, *unless* x is slowly modified over a long period by having worn out parts replaced through many operations of disassembly and re-assembly. Still our intuition is surely firm that the identity of x can be preserved through an operation of disassembly and re-assembly changing a small number of serviceable components, although these components are not at all worn and have nothing wrong with them. Surely we could disassemble our watch and replace just one of the small components within it, however serviceable or new it was, and still be able to re-assemble the same watch again. But if the first part of (T) is to allow for such possibilities, how are we to avoid a slippery slope paradox? What is to prevent a 'charade' of a too rapid series of disassembly and re-assembly operations for the wrong reasons, whatever they are?

Lowe is rightly interested in the differences between the temporal and the modal cases. As noted above, he feels that the modal analogue of (T) is quite different from (T) itself. He holds that (T) '... should be amended precisely by insisting that any large change is only consistent with an artefact's continued identity provided that it is effected through a temporal sequence of minor changes ... ' But it is pointless to say that an artefact can be assembled in a 'distant' possible world from components very different from those it has in the actual world only if there is a sequence, from the actual world to the 'distant' world, of other possible worlds in which a similar artefact is successively assembled from slightly different components. As Lowe in effect points out, this necessary condition would be pointless because such a modal sequence would always exist. Lowe's general point is perhaps summed up by Quine in the following passage about the difference between the temporal and the modal:

> The devastating difference is that the series of momentary cross sections of our real world is uniquely imposed on us, for better or for worse, whereas all manner of paths of continuous gradation from one possible world to another are free for the thinking up (W. V. Quine 'Worlds Away', in *Theories and Things*, Harvard University Press 1981, p. 127).

However, note that what Lowe calls a temporal sequence of disassembly/re-assembly episodes is not necessarily 'imposed' on us in one sense. Someone after all is responsible for these episodes. Of course, such episodes are not 'free for the thinking up' and are not always a practical possibility. Yet what *philosophical* difference does it make whether we can bring about a large change 'in one fell swoop' or in what might be called an arbitrary sequence of (perhaps very rapid) disassembly/re-assembly episodes? Lowe wants to add a necessary condition to (T), as quoted above, for when a large change will preserve an artefact's identity, but this addition to (T) is far too weak to block a temporal slippery slope paradox. We do feel strongly that some (T)-like principle is true, and yet it is unclear how to state such a principle consistently given two intuitions we apparently have. The first intuition is that an artefact can be disassembled and then re-assembled with some minor change in its components, and the second is that the artefact's identity would not be preserved through a large change in its components, whether this change took place all at once or as a result of a sequence of some kinds of disassembly/re-assembly operations which individually would be sanctioned by the first intuition. And so to repeat the conclusion of my first paper, there is a temporal slippery slope paradox, and it is sufficiently similar to the modal paradoxes, in spite of the deep differences between the temporal and the modal, for us to expect that any solution to the latter paradoxes would also apply to the former.

Sunderland Polytechnic,
Sunderland SR1 3SD.

© D. E. OVER 1986

REPLY TO OVER

By E. J. LOWE

SUPPOSE I take my watch *x* to the watchmaker to be cleaned. He disassembles it, cleans all the parts and then reassembles them, but in the process accidentally swaps one of the cogs c_1 for a very similar cog c_2 coming from another disassembled watch lying on his workbench. Almost immediately he realises his mistake, takes the watch to pieces again, removes c_2 and puts the watch together once more with c_1 in its rightful place. Should we say that at any time during this process the extraneous cog c_2 became a *part* of my watch *x*? I think *not*. It is not enough for a component to become a *part* of an existing artefact merely that it should momentarily usurp the position of one of that artefact's existing parts: rather, it has to *earn* the title by genuinely functioning in the appropriate role for a reasonable length of time. Accordingly, I cannot accept Over's assumption that a satisfactory criterion of artefact identity should admit that a very temporary displacement of a few of an artefact's parts by other similar objects amounts to a *change in its parts*. Hence I can avoid his new version of the temporal slippery slope paradox by refusing even to take the first step. That is to say, in answer to Over's question 'What is to prevent a "charade" of a too rapid series of disassembly and re-assembly operations for the wrong reasons ...?', I should say that while such a 'charade' may indeed be performed, what prevents it from threatening a paradox is the fact that even the *first* of such a series of operations (and so *a fortiori* any subsequent one) does not generate an instance of the preservation of artefact identity under a change of parts – not because artefact identity is *not* preserved here, but because parts are not in fact genuinely *changed*.

University of Durham,
50 Old Elvet,
Durham DH1 3HN.

© E. J. LOWE 1986

REPLY TO LOWE

By D. E. OVER

LOWE seems to have a reasonable doubt about the first part of (T), which he takes as implying that a new cog immediately becomes a part of a watch which is disassembled and then re-assembled with this addition. He holds that such a new cog must 'earn' its place as a part by functioning over a 'reasonable' period. One might ask here how this cog can even start to function unless it is a part of the watch. But I do accept that it may be vague when a cog is part of a watch. Still, I do not see how this point blocks a temporal slippery slope paradox. It surely does not take long for a cog to become definitely part of the watch – the reasonable period must be quite short for any particular cog. But then many of these short periods for a large number of new cogs could follow one another in a long series which did not intuitively preserve the identity of the watch. Return yet again to the example of the watchmaker. He is to repair my watch, but he takes some time over it. After I complain, he tries to give me back a watch which does not contain any of the pieces which were parts of the watch I gave him. I think I have every right to deny that the watch he tries to give me is my watch, even if he has performed many operations of disassembly and re-assembly, replacing one original piece at a time with a new piece, which he has allowed to function for a reasonable period before going on to the next operation.

Sunderland Polytechnic,
Sunderland SR1 3SD.

©D. E. OVER 1986

TARSKI – AN AMBIGUITY

By RICHARD C. JENNINGS

CONSIDER Tarski's now classic example of a T-sentence (i.e. an equivalence of the form (T)):

(C) 'Snow is white' is true if, and only if, snow is white.

There are two possible interpretations of this sentence, both of which find support in Tarski's discussion:

First, (C) can be interpreted as asserting that 'snow is white' is true if, and only if, snow is white.

> Second, (C) can be interpreted as asserting that the two sentences '"Snow is white" is true' and 'Snow is white' are equivalent.

The difference between the two interpretations is that, under the first interpretation, (C) is taken to assert the correspondence of a true sentence and a state of the world; while under the second interpretation (C) is taken to assert the equivalence of two sentences.

The first interpretation is clearly indicated by the following two remarks taken, respectively, from 'The Concept of Truth in Formalized Languages' and 'The Semantic Conception of Truth':[1]

> Amongst the manifold efforts which the construction of a correct definition of truth for the sentences of colloquial language has called forth, perhaps the most natural is the search for a *semantical definition*. By this I mean a definition which we can express in the following words:
>
> (1) *a true sentence is one which says that the state of affairs is so and so, and the state of affairs indeed is so and so*. (Tarski (1), p. 155).

and:

> *Semantics . . . deals with certain relations between expressions of a language and the objects* (or 'states of affairs') *'referred to' by those expressions*. (Tarski (2), p. 345).

The second interpretation is seen in the next two remarks, the first to do with the general form of the T-sentence and the second with the sentence (C) above:

> We ask now what is the logical relation between the two sentences '*X is true*' and '*p*'. It is clear that from the point of view of our basic conception of truth the sentences are equivalent. In other words, the following equivalence holds:
>
> (T) *X is true if, and only if, p*. (Tarski (2), p. 344)

and:

> In fact, the semantic definition of truth implies nothing regarding conditions under which a sentence like (1):
>
> (1) *snow is white*
>
> can be asserted. It implies only that, whenever we assert or reject this sentence, we must be ready to assert or reject the correlated sentence (2):

---

[1] In this paper I will refer to two of Tarski's papers: 'The Concept of Truth in Formalized Languages', in A. Tarski, *Logic, Semantics, Metamathematics*, (at the Clarendon Press, Oxford) 1956, pp. 152-278; and 'The Semantic Conception of Truth', *Philosophy and Phenomenlogical Research 4* (1943-4) 341-75. The latter is reprinted in Herbert Feigl and Wilfred Sellars, *Readings in Philosophical Analaysis* (Appleton-Century-Crofts) 1949, pp. 52-84. Page references will be to the former source. These papers will be cited as Tarski(1) and Tarski(2) respectively.

(2) *the sentence 'snow is white' is true*. (Tarski (2), p. 361)

In a recent paper I argued that Tarski's semantic conception of truth is a form of relativism ('Popper, Tarski and Relativism', ANALYSIS 43.3, June 1983, pp. 118-23). Subsequently Harvey Siegel replied ('Tarski a Relativist?', ANALYSIS 45.2, March 1985, pp. 75-6), saying that Tarski's semantic conception of truth cannot support my relativism, that:

> Tarski's theory is neutral as between alternative 'epistemological attitudes' ... regarding the existence of facts and the relation between facts and sentences. (Siegel, p. 76)

My difference with Siegel rests on the ambiguity I have just characterised: I adopt the first interpretation of T-sentences and my conclusions follow from this; Siegel adopts the second interpretation of T-sentences and his conclusions follow from that. But I will argue that the first interpretation expresses Tarski's primary concerns in developing the semantic conception of truth and that we must therefore adopt the first interpretation. Let us first see how my difference with Siegel depends on this ambiguity and then look at the reasons for adopting the first interpretation.

In 'Popper, Tarski and Relativism' I argued that Tarski's semantic conception of truth is relativistic on the following grounds. First, in developing a metalanguage that can refer to the relations between sentences and the world, it is necessary to be able to refer to the world with that metalanguage. Second, the language which Tarski introduces into the metalanguage for this purpose is taken directly from the object-language – apart from terms of syntax and of logic, Tarski does not allow any undefined terms into the metalanguage except terms of the object-language. Therefore, the metalanguage refers to the same things that the object-language does – the world of the metalanguage is relative to the world of the object-language.

For example, if the object language were the language of phlogistic chemistry, then we would find the semantic conception of truth concerned with claims such as:

> 'Phlogiston is given off during combustion' is true if, and only if, phlogiston is given off during combustion.

And in making such an assertion we are, at least according to the first interpretation of T-sentences, referring to phlogiston. The metalanguage simply adopts the ontology of the object-language, and in this sense it is relativistic.

Siegel argues that I am mistaken in asserting that Tarski's analysis supports or constitutes a form of relativism. According to Siegel:

> What correspond on the semantic conception are sentences in the object-language and sentences in the meta-language, not sentences and extra-linguistic facts. (Siegel, p. 75)

This is a clear statement of the second interpretation of T-sentences. Relativism is avoided by this interpretation because we interpret the semantic conception of truth as concerned with the relations between sentences (sentences in the metalanguage and sentences in the object-language) and thus we are able to avoid any of the epistemological or ontological questions that can be raised about talk of the world. That is, under the first interpretation, when the metalanguage refers to the world, we can raise epistemological and ontological questions about such reference. But when we confine our attention to relations between sentences, such questions have no place and, as Siegel says, Tarski's theory is neutral between alternative epistemological or ontological views.

Now, in 'Popper, Tarski and Relativism' I did not distinguish these two interpretations of T-sentences and, because of this, my argument was flawed. On the basis of the second interpretation, I argued that Popper could not use the semantic conception of truth to support this view of objective or absolute truth; and on the basis of the first interpretation I argued that Tarski's semantic conception of truth was relativistic. It was Siegel's reply that drew my attention to this ambiguity. What I will now argue is that if we are to be true to Tarski's original intentions in developing the semantic conception of truth, we must adopt the first interpretation of T-sentences. This results in Tarski's analysis supporting a form of relativism, and in the claim that Popper's theory of truth, insofar as it is based on this analysis, is relativist as well.

The question whether Tarski intended the first interpretation or the second boils down to the queston whether Tarski aimed to provide an analysis of the correspondence theory of truth, that is, whether he intended to explicate the relation between true sentences and states of the world. And there are strong reasons to think this was his intention. In 'The Concept of Truth in Formalized Languages' Tarski asserts that he will be:

> ... concerned exclusively with grasping the intentions which are contained in the so-called *classical* conception of truth ('true – corresponding with reality') ... (Tarski(1), p. 153)

In 'The Semantic Conception of Truth' he says he would like his definition of truth to do justice to the classical conception of truth which he characterises in, among other ways, the following way:

> *The truth of a sentence consists in its agreement with (or correspondence to) reality.* (Tarski(2), p. 343)

Further on, concerning the sentence 'snow is white', Tarski states that:

> ... if we base ourselves on the classical conception of truth, we shall say that the sentence is true if snow is white, and that it is false if snow is not white. Thus, if the definition of truth is to conform to our conception, it must imply the following equivalence:

> *The sentence "snow is white" is true if, and only if, snow is white.* (Tarski(2), p. 343)

Thus, the T-sentences are intended to provide a measure of the adequacy with which the semantic conception of truth does justice to the correspondence theory of truth.

I conclude that Tarski's semantic conception of truth is intended to explicate the correspondence theory of truth, that only the first interpretation of T-sentences is acceptable for this purpose, and, finally, that, because of this, Tarski's analysis does actually support or constitute a form of relativism.

University of Edinburgh,
David Hume Tower,
George Square, Edinburgh EH8 9JX

© RICHARD C. JENNINGS 1986

REPLY TO GARRETT

By HAROLD W. NOONAN

IN my article 'Wiggins, artefact identity and "best candidate" theories' (ANALYSIS 45.1, January 1985, pp. 4–8) I criticized the 'best candidate' approach to identity over time, that is, the approach according to which whether a later individual, y, is identical with an earlier individual, x, may be determined merely by whether, at the later time, there are any *better candidates* than y around for identity with x.

In his interesting reply 'Noonan, "best candidate" theories and the Ship of Theseus' (ANALYSIS 45.4, October 1985, pp. 212–15) B. J. Garrett comes to the defence of the best candidate theorist. In what follows I explain why I am still unconvinced.

Much of my discussion centred on the Ship of Theseus problem. The three relevant situations are pictured below:

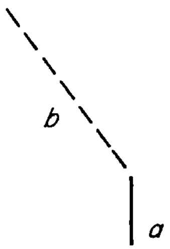


Figure 1

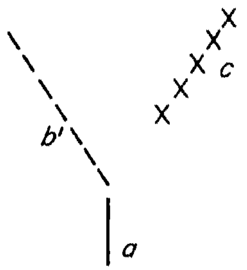


Figure 2

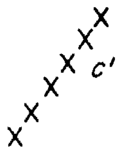


Figure 3

In all three situations, '*a*' denotes the Ship of Theseus. In situation 1, '*b*' denotes the continuously repaired ship. In situation 2 (the one described by Hobbes), 'b'' denotes the continuously repaired ship and '*c*' denotes the plank-hoarder's ship. In situation 3 the ship reconstituted from *a*'s original planks is denoted by 'c''. Given that we regard spatio-temporal continuity (under a sortal) as the most important criterion of artefact identity, outweighing the identity of original parts criterion, but acknowledge that an artefact can survive dismantlement and reassembly, the most natural description of the above situations is as follows: in situation 1, $a = b$; in situation 2, $a = b'$ and $a \neq c$, and in situation 3, $a = c'$. This common-sense description coincides with that of the best candidate theorist.

However, I argued that this description of the three situations was unacceptable. For to accept it committed one either to denying the necessity of identity or to endorsing a proposition I characterized as 'self-evidently absurd'.

Now according to Garrett there are at least two rejoinders the best candidate theorist can make to the criticisms of his position I put forward: one is to deny the transitivity of identity, which was an implicit assumption of my reasoning; the second is to claim that the proposition which I assert that the best candidate theorist must accept if he accepts the necessity of identity is not, as I claim, self-evidently absurd, because in accepting it he is not accepting anything important.

I shall consider these suggestions in turn.

Actually all that needs to be said about Garrett's first suggestion is that the transitivity of identity is an undeniable fact, and that it is no more open to the best candidate theorist to deny it than it is open to him to deny the Principle of Non-Contradiction.

However, Garrett produces an argument in support of his claim that one can reasonably consider denying the transitivity of identity which is worth further scrutiny.

Garrett imagines an artefact, a table, which is gradually transformed into a chair. He imagines that before any change is made the table is dubbed '*A*', after a change too small to make it plausible to claim that the original artefact no longer exists the artefact present is dubbed '*B*', after a change of a similarly limited type the artefact present is dubbed '*C*' ... and so on until eventually the artefact present is a chair, which is dubbed '*Z*'. For future reference let us suppose that throughout this period the artefact (artefacts) in question is (are) located in a place *L*, and let us call the times at which the various dubbings occur, 'moments of ostension' and denote them by '*tA*', '*tB*' ... '*tZ*'.

Now Garrett asserts, plausibly enough, that no artefact is located in *L* both at the moment of ostension *tA* and at the moment of ostension *tZ*, so that '$A = Z$' is definitely false. But, he asserts, each of '*A* = *B*' '*B* = *C*', . . ., '*Y* = *Z*' is true, whence he concludes that identity is non-transitive.

Given that identity is transitive, and given that we accept Garrett's claim that no single artefact is located in *L* both at *tA* and at *tZ*, as I think we should (for otherwise what kind of (gradual) Protean change *could* we disallow?), we must contrapose: at least one of '*A* = *B*', '*B* = *C*', . . . must fail to be true. But Garrett is surely right that it is implausible to suppose that any one of these identity claims is definitely *false*. Whence it follows that one of them must be neither true nor false, i.e. indeterminate in truth-value. However, a statement of identity can be indeterminate in truth-value only if one or both of its terms lacks a determinate reference. (For the argument for this see G. Evans 'Can there be vague objects', ANALYSIS 38.4, October 1978 and my 'Vague Objects', ANALYSIS 42.1, January 1982 and 'Indefinite Identity: A Reply to Broome', ANALYSIS 44.3, June 1984.) At least one of '*A*', '*B*', . . . must then lack a determinate reference, and so, at at least one moment of ostension, i.e. at *tA* or *tB* or . . . *tZ*, there must be present in *L* two distinct objects, each of which is, under at least one acceptable sharpening of the vague term 'artefact', an artefact.

But we can go further. Though, given the transitivity of identity, we cannot accept that each of the series of identity claims '*A* = *B*', . . . is definitely true, there is a weaker series of claims, entailed by these, which we can accept, namely the series asserting of each pair of successive moments of ostention that a single artefact is located in *L* at both these moments, i.e. the series of claims (1) 'There is a single artefact in *L* both at *tA* and *tB*', (2) 'There is a single artefact in *L* both at *tB* and *tC*', . . . (25) 'There is a single artefact in *L* both at *tY* and *tZ*'. But if these claims are all definitely true, and it is also definitely true that *no* single artefact is located in *L* both at *tA* and *tZ*, it follows that it is definitely true that *at some moment of ostension or other* during the relevant period *two* artefacts (i.e. artefacts with distinct histories) are simultaneously present in *L*. (For there is an artefact in *L* at *tA*. Now suppose that this does not cease to exist until after *tY*, then at *tY* both it and an artefact which by claim (25) exists both at *tY* and at *tZ* are simultaneously present in *L*. The same argument applies whenever after *tA* we suppose the artefact we start out with ceases to exist.)

To summarize: if we maintain, as we must, that identity is transitive, we cannot describe the situation imagined by Garrett, as he suggests we should, by the series of identity claims '*A* = *B*', . . . '*Y* = *Z*', together with the non-identity claim '*A* ≠ *Z*'. However, we can accept Garrett's implicit claim (a) that a single artefact is in *L* at any two successive moments of ostension but that no single artefact is in *L* both at *tA* and *tZ*. But if we do so we must accept

(b) that it is definitely true that at some moment of ostension or other two artefacts are simultaneously located in *L*.

This latter proposition may or may not seem quite bizarre. But whether it does or not is quite irrelevant. My claim is just that given the transitivity of identity (b) is entailed by (a) and so, since the transitivity of identity is undeniable, to accept (a) one must accept (b). In other words I have not been trying, in the preceding discussion, to *defend* the transitivity of identity, which is no more possible than to defend the Principle of Non-Contradiction; I have simply been considering what it is open to us to say of the interesting situation described by Garrett given that we acknowledge that identity is transitive.

I now turn to Garrett's second proposed rejoinder to my criticism of the 'best candidate' theory.

I argued that if the best candidate theorist did not reject the necessity of identity (or reject transitivity) he was committed to saying of *c* in situation 2 that, but for the existence there of *b′*, it would not have existed at all, and this I claimed was self-evidently absurd. Garrett accepts that this proposition is a consequence of the best candidate theory, as applied to the case of the Ship of Theseus, given the necessity of identity, but disputes its absurdity. For he says

> although it is true that, but for the existence of *b′*, *c* could not have existed, this fact is not important given that a ship, materially indiscernible from *c*, would then have existed, which stands in the same material relation to the Ship of Thesus as *c* does in situation 2 ... if the 'best candidate' theorist takes the Parfitian view that identity, and hence existence, is not 'what matters', the fact that, but for the existence of *b′*, *c* would not have existed at all, ceases to be a significant fact, and hence ceases to be an absurd one.

I shall come back to this reply in a moment. First, however, I need to acknowledge an error in my earlier article. For, in fact, it is not a consequence of the best candidate theory, together with the necessity of identity, that *c* in situation 2 would not have existed if *b′* had not. In situations in which one candidate is superior to the rest, according to the best candidate theory, the superior candidate *is* the original. But situation 2 is of this kind. Hence *b′* is *a*, i.e. the Ship of Theseus, and so *b′*, i.e. *a*, does exist in situation 3 (where no repair and replacement work takes place) even if, as I argued the best candidate theorist had to accept, *c* does not. But then the existence of situation 3 does not, as I assumed, commit the best candidate theorist to the claim that if *b′* had not existed *c* would not have done so either. In fact there is no reason why the best candidate theorist should not regard this as simply false.

However, what the best candidate theorist *is* committed to saying is that a situation differing from situation 2 in a mere Cambridge way with respect to what happens in the location of *c* – namely, by not containing the events there contained in the history of *b*′ (*a*) after repair and replacement work begins – may be a situation from which *c* is absent. But this seems to me quite as absurd as the proposition I originally claimed the best candidate theorist to be committed to asserting. In addition the best candidate theorist is committed to saying of any pair of *equally good candidates* for identity with an earlier individual (than which there are no better), that it is true of each of the pair that, if it had not existed, nor would the other, and this too seems to me self-evidently absurd.

But Garrett would disagree, for the reasons outlined in the quotation I gave, so I must now consider his argument. As I understand it, it amounts to the claim that because the truth-value of these propositions, which I claim to be self-evidently absurd, is of no practical importance, they cannot be self-evidently absurd.

However, why on earth should this be so? Why can it not be the case that something is from a practical point of view of no significance whatsoever, and yet the truth about it still be obvious? Surely this is very often the case.

I could leave my reply to Garrett at that, except that he refers to Parfit's thesis that identity is not what matters in survival and suggests that it is in some way the crucial premiss of his response. This may lead the reader to think that something more profound is going on in Garrett's argument than the diagnosis just given uncovers. However, I do not think that this can be so. In fact, I shall argue, Parfit's thesis is quite irrelevant to the issue between myself and Garrett.

To see that this is so it needs to be understood what Parfit's thesis is. As I understand it, it is the claim that, contrary to what we are all naturally inclined to believe, we do *not* have a basic and non-derivative concern for our own future existence. What is of fundamental importance to us is that there be in the future people related by certain relations of psychological continuity and connectedness to ourselves as we are now (let us call such future people 'Parfitian survivors'). Now, in the actual world, at the present time, the only way I can secure that I have a Parfitian survivor tomorrow is to ensure that I myself am around tomorrow – that is, that I still *exist* tomorrow, that one of the people alive tomorrow is *identical* with me. Nevertheless, my having a Parfitian survivor tomorrow does not *entail* that I exist tomorrow and, in certain conceivable, or possibly actual but merely future, circumstances, it will be possible to ensure a Parfitian survivor for oneself without ensuring one's own future existence. Parfit's thesis that identity is not what matters in survival is then the thesis that, given one's fundamental desires and concerns, one would have no reason, in such a situation, for preferring a future in which one was present oneself to one in which one merely had Parfitian survivors.

Now this is an extremely surprising claim, but Parfit has an argument for it, to the effect that the explanation of our intuitions about fission cases requires its acceptance. In fact, I think that this is not so. But even if this thesis were wholly uncontentious I do not see that it would be relevant to the issue between Garrett and myself.

Garrett thinks otherwise because he takes his argument against me to be the same as an argument of Parfit's against an objection to Reductionist theories of personal identity. Reductionist theories of personal identity, of the type favoured by Parfit, are of the 'best candidate' variety, and thus conflict with the Only x and y principle, that is the principle, to state it in the rough way favoured by Parfit (*Reasons and Persons*, Oxford, 1984, p. 267), that whether a later person *y* and an earlier person *x* are the same person can depend only on the *intrinsic* relations between *y* and *x*. But in *Reasons and Persons* Parfit argues that the plausibility the Only *x* and *y* principle undoubtedly has can be explained, given this thesis that identity is not what matters in survival, in a way that does not require us to acknowledge its correctness. According to Parfit the reason why we find the Only *x* and *y* principle plausible is that we mistakenly believe that identity *is* the relation that matters in survival. Because we do this we illegitimately ascribe properties of the relations that do matter to identity. Now it *is* the case that whether the relations that matter in survival (psychological continuity and connectedness) hold between an earlier person and a later one cannot be determined extrinsically, and so we find it plausible to say the same of identity, that is, to accept the Only *x* and *y* principle.

This is, I think, a sound argument, given its premiss. But, of course, it can only explain our finding the Only *x* and *y* principle plausible as a constraint on our concept of *personal* identity. It has no power at all to explain why we find it plausbile more generally, for ships, for example. For the thesis that identity is not what matters in survival is not a thesis about identity in general, but specifically a thesis about the nature of our concern for our own future existence, namely that it is a derivative concern – a concern for a means to an end rather than for an end in itself (the end being, of course, the existence of Parfitian survivors).

But then, as I claimed earlier, this thesis cannot possibly be relevant to the dispute between Garrett and myself, which concerned the absurdity or otherwise of a proposition stating that the *ship c* would not have existed in certain specified circumstances. And, consequently, Garrett's argument, despite appearances, can be no more profound than the original diagnosis I gave of it suggested.

In a nutshell, what Garrett fails to see is that Parfit's assertion that identity is not what matters in survival is a *slogan*, which takes its meaning from the context of its surrounding arguments in

Parfit's text. But when this context is taken into account the slogan turns out, despite its apparent generality, to express a claim purely about personal identity. Taken out of context the slogan has no definite meaning at all.

University of Birmingham,
P.O. Box 363, Birmingham B15 2TT

© HAROLD W. NOONAN 1986

GRAYLING, FEYERABEND AND THE CONSTANCY OF SENSE

By JOHN SHAND

IN his recent book, *The Refutation of Scepticism* (London: Duckworth, 1985), Anthony Grayling argues against Feyerabend's contention that radical change in the meanings of theoretical terms destroys continuity of reference between these terms, and since reference is tied to sense, that this entails incommensurability between theories. Whatever we may think of Feyerabend's arguments for this view, I wish to contend that he cannot be wrong for the specific reason given by Grayling in the example of temperature, and of the different calibrated systems we have for measuring it (*ibid*: pp. 56–9). This is important because the specific failure of Grayling's argument is indicative of the problems besetting any enterprise which aims to show that there are necessary conditions to all experiences, as Grayling wishes to. Suggested 'necessary conditions' turn out to be either vacuous, or dubious in their necessity, and not indispensable to our conception of the world. The temperature example is a case of the dubious variety.

We can note that, if terms x and y do not overlap in meaning, then no sense can be given to an improvement in explanation in moving from x to y. In general this can be granted, although the account cannot be as simple as this because it ignores the fact that x and y cannot be considered as isolated theoretical terms, but only as part of a scheme. While it is a requirement, in comparing x and y, that they have an overlap of sense, it does not follow that the area constituting the overlap is somehow necessary and immutable, so that all terms of *types* x and y must overlap in this one sense, if not in any other. All that is entailed is that there is an overlap of some kind. Grayling tries to tie down this overlap in the case of temperature and the different ways we measure it to '... an ability to perceive sensible differences in heat – "hotter" and "colder" form the scale which counts, and that must be invariant with respect to sophisticated means of dealing with temperature' (p. 59). He may even regard the ability to perceive sensible relative, sufficiently diverse temperature differences ('hotter' and 'colder') as a necessary feature of experience, and indeed he must if Feyerabend's case is to be refuted. But Grayling's argument is refuted by the case of our sensible experience of frozen carbon dioxide. Touching frozen carbon dioxide, particularly when blindfolded, induces a sensation of intense heat, and the person having the experience will say he is certain he has been burnt. So the unsophisticated discrimination of 'hotter' and 'colder' is immediately dislocated, and is certainly denied its basic, invariant or logically prior status of '... the appearances which the sophistications must save' (p. 59). The transfer of a hand from immersion in hot water to hot water, and then cold water to hot water, causes similar misjudgment. Our measured temperature overrides the appearance. We rely on our ability to measure temperature, not on what we perceive sensibly, to tell us that frozen carbon dioxide is colder than a luke-warm bath. We do this, because to assert that frozen carbon dioxide is intensely hot (hotter than skin temperature, say) would completely abandon great portions of our theoretical framework concerning the behaviour of substances with respect to changes in the kinetic energy of their molecules. It is only for a minute range of measured temperatures that sensible perception has any relevance at all. Experience may be used by analogy as a heuristic and imaginative device (cf. Quantum mechanics). But the difference between 1 million degrees and 10 million degrees centigrade has nothing to do with our 'ability to perceive sensible differences', or any direct experience of the difference, but everything to do with our theories of heat and temperature.

King's College,
Cambridge

©John Shand 1986

THE POSSIBILITY OF TRANSPARENT WHITE

By David H. Sanford

C. L. HARDIN begins 'A Transparent Case for Subjectivism?' (Analysis 45.2, March 1985, pp. 117-19) by quoting what Colin McGinn calls a tempting objectivist reduction. The reduction purports to explain the impossiblity of transparent white:

> For a surface to be transparent is for it to let through the incident light, whereas for it to be white is for it to reflect all of the light; so transparent

> white is impossible because a surface cannot simultaneously interact in both these ways with the light incident upon it (*The Subjective View*, Oxford, Clarendon Press, 1983, p. 36).

A perfect mirror that reflects all incident light, Hardin points out, is not white. White surfaces reflect diffusely. I would add that although a surface of a transparent object lets through incident light, objects rather than surfaces are transparent. Light passing through one side keeps going through a transparent object to emerge on the other side. The impossibility of a single surface both reflecting and transmitting most of the incident light does nothing to show that something cannot be both transparent and white, for the relevant reflecting and transmitting surfaces are distinct. Consider a sheet of glass or plastic with opposite surfaces *A* and *B*. Most of the light incident on *A* is diffusely reflected. Most of the light incident on surface *B* goes straight through and out the other side, surface *A*. How this white transparent sheet looks from side *A* depends on the difference between the levels of illumination on the two sides. One can see through from the darker to the brighter side. Gauze hanging between the audience and the back part of the stage presents a similar phenomenon.

Some white fabrics are transparent. They can be seen through because of spaces between the threads. They are white because they are made of white threads. Arguments to show the impossibility of things both white and transparent should be suspect, since there actually are things both white and transparent.

But something both white and transparent is not thereby *transparent white* as the phrase is understood in this context. White gauze does not whiten the light that passes through it. Things seen through white gauze do not look whiter. Transparent red glass, in contrast, reddens the light that passes through it so that some things viewed through it look redder. If there were transparent white glass, it would whiten light so that some things seen through it looked whiter. But (1) transparent glass never whitens light so that things viewed through it look whiter. Hardin mentions an interesting transposition about mirrors: (2) mirrors also never whiten light so that things viewed by their means look whiter.

Truths (1) and (2) are contingent. Some small tales of technology fiction show how they might be false; science fiction stories about objects and events that violate physical law will not be needed. There is nothing physically, conceptually, or phenomenologically impossible about white mirrors or transparent white.

D. K. Buckner restates one of Hardin's points as follows: 'all white objects are non-selective diffusing reflectors' ('Transparently False: A Reply to Hardin', ANALYSIS 46.2, March 1986, p. 86). We must first remember that something can be white without reflecting light in a certain way. Light itself can be white, as can some emitters of white light such as white hot iron bars, light bulbs, the sun, and white stars. 'What we see as luminous we don't see as grey. But we can certainly see it as white' (Ludwig Wittgenstein, *Remarks on Colour*, III. 226). While reflective white is at one end of an achromatic continuum with black at the other end and degrees of grey in between, luminous whites are brighter and dimmer, but not more or less grey. Brown is also never luminous (cf. *Remarks*, III, 215), but many colours in addition to white can be either luminous or reflective. My descriptions of transparent white and white mirrors combined some properties of light emitters with some properties of ordinary mirrors and transparency.

A bullhorn is a hand-held combined microphone and loudspeaker. When you speak softly into one end, the louder sound coming out the other end is not just louder: its character depends on the character of the sound coming in. It is your voice amplified. A bullhorn uses electric current, usually provided by batteries. Like any sound-amplifying system, it has its limits: there will be some sound too loud for it to amplify further.

The visual analogue of a bullhorn is a bullwindow. When dim light enters one side, brighter light is emitted from the other side. The light coming out is not just brighter: its character depends on the character of the entering light. A bullwindow is a light amplifier. (Bullwindows, like bullhorns, use electric current and have their limits.) Given any light path passing through a standard bullwindow, when light of any mixture of frequencies travels along it and enters the receiving side of the window, then light of the same mixture of frequencies but with higher intensity emerges along that path from the other side of the window. Dim things look brighter and bright things look brighter still. Something perfectly black, which reflects no light at all, would continue to look perfectly black through an ideal bullwindow. (Less than ideal bullwindows emit some 'noise' light even when no light at all impinges on the other side.)

Ordinary transparent windows of coloured glass colour light by selective absorption. Light of some frequencies is absorbed more than light of other frequencies. Coloured bullwindows colour light by selective amplification. Light of some frequencies is intensified more than light of other frequencies. When viewed through ordinary transparent red glass, something green looks less green and darker. When viewed through one kind of transparent red bullwindow, something green looks less green and brighter. In both cases, the green looks more achromatic, not more red. In one case, the green frequencies are absorbed more than some others. In the other case, some other frequencies are amplified more than the green frequencies. Bullwindows of this kind make red things and grey things look redder and brighter, yellow things look brighter and more orange, blue things look brighter and more magenta, and green things look brighter and more achromatic. Some green things look whiter.

A transparent green bullwindow similarly makes fairly bright red things look pinker, that is, whiter. Red things look brighter and more achromatic.

A transparent white bullwindow makes everything look brighter and everything not already achromatic look more achromatic. Less intense light of a certain frequency is more amplified than more intense light of another frequency entering at the same time along the same path. Everything looks brighter and every hue becomes less saturated. Not everything looks whiter: the dark greys, greens, blues, violets, and purples all look less dark, but they do not move close enough to white to look whiter. In general, if it is possible for something of certain colour to look whiter, it can look whiter because it is viewed through a transparent white bullwindow. Transparent white bullwindows whiten light.

Bullmirrors amplify light just as bullwindows do except the light path changes direction as it would with an ordinary mirror. Bullmirrors reflect more light than they receive (within limits). White bullmirrors amplify less intense frequencies more than the more intense. The character of the light reflected by the white bullmirror can be just the same as that of light transmitted by a transparent white bullwindow from the same scene. White bullmirrors whiten light.

Wittgenstein assumes throughout *Remarks on Colour* that any glass not perfectly clear and colourless absorbs some of the light passing through it. Hardin says that a coloured transparency 'will necessarily darken some colour' (p. 119). I question the necessity. Just as the absence of light amplifiers helps explain the absence of transparent white, the possibility of light amplifiers helps explain the possibility of transparent white.

Duke University,
Durham, NC 27708, USA

©DAVID H. SANFORD 1986

NOONAN ON TEMPORAL PARTS

By GRAHAM SPINKS

IN his article 'A Note on Temporal Parts' (ANALYSIS 45.3, June 1985) Harold Noonan argues that

> *if* enduring objects have temporal parts then it is reasonable to say that (some of) the temporal parts of an enduring object of some sort S are themselves objects of sort S. If persons have temporal parts it is reasonable to say that some of these are persons, if tables have temporal parts it is reasonable to say that some of these are tables, and so on.

Noonan's strategy is not to provide a positive argument for his claim but rather to consider and dismiss two commonly made objections. Below I shall sketch a further objection which I believe requires an answer.

In order to sidestep issues specifically concerning *personal* identity let us restrict our discussion to the less perplexing sortal concept 'table'. Consider the following:

(a) temporal parts of tables (TPTs), which endure over a sufficiently long period of time t, are themselves tables;
(b) during any period of time longer than t, in a room containing at least one table, many TPTs having a lifespan of longer than t will come into and pass out of existence;
(c) during any period of time longer than t, in a room containing at least one table, many tables will come into and pass out of existence.

Now, presumably, the reason why Noonan hesitates to claim that *all* temporal parts of objects of sort S are themselves of sort S is that some of them are too fleeting. (The grounds for such a reservation are obviously clearer in the case of persons than in that of tables). (a) is, therefore, intended as a clarification of Noonan's basic claim insofar as it affects tables. Anyone prepared to countenance TPTs at all should, surely, be happy to agree to (b). However, (c) which follows from the conjunction of (a) and (b) is false. It is not the case that tables are constantly coming into and passing out of existence in any room containing a table for a period of time longer than t (however long that may be). To claim otherwise would be to thwart usage and, thereby, to contravene meaning. Thus the criteria for TPTs having a lifespan of longer than t coming into and passing out of existence are very much more liberal than those applying to tables. Whenever a table comes into or passes out of existence a number of TPTs having a lifespan of longer than t will do so simultaneously, but not vice-versa.

Generalising from this representative example we may conclude that temporal parts of enduring objects of some sort S cannot themselves be of sort S since the coming into and passing out of existence of a temporal part of an object (no matter how long lived) of sort S is not a sufficient condition of the coming into or passing out of existence of an object of sort S.

Darwin College,
Cambridge, CB3 9EU.

© GRAHAM SPINKS 1986

MORE PARADOXICAL EPISTEMICS

By MARTIN HOLLIS

A THINKS of a number and whispers it privately to *C*. *B* does the same. *C* declares, 'You have chosen different positive whole numbers. Neither can determine whose is the greater.' That could apparently be true – or so it seems to *A* who happens to have picked 157. But then *A* reflects that it implies that neither has picked 1 and hence, granted this much, that neither has picked 2. Anyone with 2, who realises that no one picked 1, knows that 2 is the smaller number. The 'Paradoxical Train of Thought', which I described in ANALYSIS 44.4 (October 1984), grinds into motion, heading for absurdity. *A* finds that either he cannot have picked 157 (which he did) or *C*'s declaration is false, thereby generating further trouble.

Doris Olin in ANALYSIS 46.1 (January 1986) tries to stop the train at the first turn of its wheels. Michael Kinghan in the same issue allows the first turn and tries to stop it at the second. I shall reply briefly to both and then suggest that the paradox is wider than the Surprise Test, to which it is clearly kindred, by proposing further variants.

Olin denies that, if *B* had picked 1, he could infer that *A*'s number was greater, because he would have as much reason to doubt *C*'s declaration. (She is inclined to focus on the clause where *C* says that the numbers chosen are different. But this can be countered by having *A* and *B* draw their numbers from a hat, which they know to contain tickets numbered 1 to *n*, or – at some cost to the neatness of the prose – by dropping the clause and letting *C* declare only that neither can determine whether the other's number is at least equal to his own.) The catch, she says, is that the train of argument requires that one (or both) of them at some point would have to be justified in believing 'his number is greater than mine and I am not now justified in believing it'. Quite so. But why conclude that 'the paradoxical train of thought is brought to a halt'? On the contrary, she has thus shown *C*'s perfectly plausible declaration to be implictly paradoxical after all.

Kinghan would presumably agree, since he lets the train at least leave the station. His initial complaint is that *A* cannot use the fact that he chose 157 as a proof that *C*'s declaration was true. Quite right. But the question, of course, is what follows *if* *C*'s declaration is true. Kinghan grants the point, when trying to salvage 'a semblance of paradox' by examining the move from 'neither chose 1' to 'neither chose 2'. Here he claims to detect a 'fallacy of epistemic distribution', occurring if one argues from '*N* knows that *p* implies *q*' to 'if *p* then *N* knows that *q*'. It crops up, he says, when *A* reasons that *B* cannot have chosen an integer all of whose predecessors are eliminated on the supposition that

C's declaration is true. But why? *A* is trying to work out what *C*'s declaration implies. He knows (or claims to) that if it is true, then *B* did not choose 1 and can infer that *A* did not choose 1. Hence *A* knows that *A* and *B* can both infer from *C*'s declaration that no one chose 1 and hence that no one chose 2. Whatever else may be wrong, the epistemic operators seem to me to be behaving impeccably.

While awaiting further objections, I would like to suggest that the relation between this paradox and the 'Surprise Test' is that the latter belongs to a larger group, which raises a general question about epistemic inference. Here is a generalization from two to *n* ensnared persons and then a fresh quirk of the epistemic calculus.

I Sheikh Inbed the Terrible, growing old, summoned his scores of wives and said, 'When I am gone, my thousands of rubies will be shared among you in a way which I have already decreed. Each of you will receive at least one. Each may wonder whether there is at least one other wife who has received less. No one, on being told only how many she herself has been allotted, will be able to decide.'

II To his two oldest wives he added, 'I am leaving my vellum copies of the *Tractatus* to be shared between you, each receiving at least one. There are in total either 99 or 100. Neither, on being told her own share, will be able to infer exactly how many the other received.' Whereat Fatima remarked, 'Then neither can receive 99. So anyone receiving 1 could place the other with 98. So anyone receiving 98 could place the other with 2. So anyone receiving 2 could place the other with 97. This reasoning continues until all possibilities have been eliminated. Inbed, you are an old fool.'

School of Economic and Social Studies,
University of East Anglia,
Norwich NR4 7TJ.

© MARTIN HOLLIS 1986

REPLY TO LOWE

By HAROLD W. NOONAN

E. J. LOWE ('Noonan on Naming and Predicating,' ANALYSIS 46.3, June 1986, p. 159) is baffled by my maintaining (H. W. Noonan, 'Relative Identity: a Reconsideration,' ANALYSIS 46.1, January 1986, pp. 6-10) that something which is incapable of being named by a name *for* an *f* may yet be predicatively an *f*, where '*f*' is a sortal predicate, e.g, that something which is incapable of being named by a name *for* a cat may yet be predicatively a cat.

In order to make clear why I think this can be so I shall refer to an instructive example of R. M. Chisholm's ('Reply to Shoemaker' in *Perception and Personal Identity*, ed. N. S. Care and R. Grimm, The Press of Case Western Reserve University, Cleveland, Ohio, 1969)

Chisholm imagines a ship, the USS *South Dakota*, which undergoes a piecemeal replacement of parts as pictured below:

VZ	iZ	ij	ij	ij	ij
WY	WY	WY	kY	kl	kl
X	X	X	X	X	m
(*a*)	(*b*)	(*c*)	(*d*)	(*e*)	(*f*)

a is the collection of planks which constitutes the ship on day one, *b* on day two and so on. There are then six non-identical plank collections, and each is of course non-identical with the USS *South Dakota*. But, Chisholm points out, Webster's definition of a 'ship' is 'a structure used for transportation in water', and each of *a* to *f*, as well as the USS *South Dakota* satisfies *this* definition. But of course, there is only one ship present in the situation diagrammed. So, as Chisholm says, we are confronted with a philosophical problem. How should we solve it?

There are three possible lines of solution:

(1) One can just say that Webster is wrong, and the correct definition of 'ship' applies to none of *a* to *f*. The problem then disappears.

(2) One can say that what the puzzle shows is that it is a mistake to suppose (as philosophers are prone to) that in everyday life counting is always by identity, i.e. that x and y are counted as one just in case x = y. In fact, in counting ships we count, as the puzzle reveals, by a weaker equivalence relation *R*. This equivalence relation obtains between each of *a* to *f* and the rest, and between each of these and the USS *South Dakota*. Consequently we are speaking correctly when we describe the situation as one in which there is just one ship even though the situation contains seven distinct objects each of which qualifies as a ship. But, of course, in counting ships the relation we count by can be none other than the one we express by 'is the same ship as'. This, then, must be the relevant relation *R*. Thus 'is the same ship as' is an expression for a relation which does not ensure the indiscernibility of its terms, i.e. as Geach would put it, it is an expression for a relative identity relation, or, as Chisholm would put it, it is an expression for identity only

in a 'loose and popular sense' not in a 'strict and philosophical sense'. (Chisholm's thesis that personal identity is identity in a strict and philosophical sense and bodily identity is identity only in a loose and popular sense is then just the thesis, to put it in Geachian terms, that personal identity is absolute identity and bodily identity relative identity.)

(3) Finally, one can say that the solution to the puzzle lies in recognizing that in counting ships not everything that qualifies as a ship counts. 'There is just one ship in place *p*' means 'some ship is in place *p* and every ship in place *p* is identical with that one'. Thus the only entities to be counted when counting ships are those which fall within the range of the natural language quantifying expressions 'some ship' and 'every ship'. But it is only if 'some ship is F' is equivalent to 'something is a ship and is F' and 'every ship is F' is equivalent to 'everything, if it is a ship, is F' that these quantifying expressions must be taken to range over everything which qualifies as a ship. A solution to the problem can thus be found in denying these equivalences, and maintaining that, of the seven items in the situation which qualify as ships, only one – the USS *South Dakota* – falls within the range of 'some ship' and 'every ship'.

In the article Lowe criticizes I argued that there could be no good reason for preferring type (2) solutions to problems of this sort to type (3) solutions, since the distinction between restricted and unrestricted quantification upon which the type (3) solution relies is something to which a proponent of a type (2) solution is anyway committed. For, even though, in Chisholm's example, *a* is a ship, according to a defender of a type (2) solution, he must acknowledge that if, for example, *a* ceases to exist on day two, it does not follow that *some ship* then ceases to exist. On grounds of economy, then, I suggested that type (3) solutions to problems of the sort Chisholm describes must always be preferable to type (2) solutions.

Lowe, however, would reject both these lines of solution in favour of solution (1). But I do not see that it is the business of the philosopher, *qua* philosopher, to tell the lexicographer his job. As Chisholm points out, Webster's definition of 'ship' does seem deficient in certain respects: surely something can qualify as a ship even if it never leaves port, and so is never used for transportation in water, and one would be doubtful about calling a structure used for transportation in water a ship if it could only stay afloat while in motion and sank like a stone when at rest. But if the definition were revised to remedy these defects, the philosophical problem would still remain; so, as Chisholm says, we need not concern ourselves with them. Lowe might say that any definition of 'ship' which allows the existence of Chisholm's problem *must* be defective, but that is just prejudice.

But, of course, the USS *South Dakota* is, and *a* is not, capable of being named by a name *for* a ship. So, Lowe will ask, in what sense *can* it be true that each is predicatively a ship? But the answer to this question has already been given: it is true in the sense given in the dictionary – each is a structure used for transportation in water.

But then what is the distinction between those entities which are, and those which are not, capable of being named by a name *for* a ship? Answer: something is capable of being named by a name *for* a ship just in case its possession of a property P is a sufficient condition of the truth of 'some ship has property P' and a necessary condition of the truth of 'every ship has property P', for 'some (every) ship has property P' is true just in case 'N has property P' is true for some (every) interpretation of 'N' as a name *for* a ship.

In conclusion, I wish to add the following points.

I have used the locution 'is predicatively a ship' – in implicit contrast, of course, with the locution 'is constitutively a ship'. What then do I take to be the role of the 'is' of predication? Here I follow Frege: the 'is' of predication (unlike the 'is' of constitution, if it exists) is just a grammatical device – a mere fragment of a predicate which expresses no property or relation by itself. But then what is the connection between the 'is' of predication and the 'is' of identity? Simply this: x is predicatively f just in case it is identical with something which is predicatively f, i.e. fx iff $(\exists y)(x = y \,\&\, fy)$. As for the 'is' of constitution: I do not wish to deny that such a sense of 'is' exists. But the equivalence of for example, '*a* constitutes a ship' and '*a* is a ship' is no evidence for it (one might as well cite the equivalence of 'Socrates killed Socrates' and 'Socrates killed himself' as evidence that 'killed himself' can mean 'killed Socrates'), and neither a defender of Leibniz's Law, nor, if I am right about the viability of type (3) solutions to problems of the sort Chisholm describes, an opponent of relative identity, *needs* to accept the existence of an 'is' of constitution.

In my previous article I illustrated my argument by an appeal to Geach's example of the cat on the mat. In this reply I have chosen to use Chisholm's example instead. I have done so in order to take the opportunity to draw attention to the great similarity, indeed partial identity, between the views of Chisholm and Geach on these topics, which has gone largely unnoticed. Of course, 'ship', unlike 'cat', is not a natural kind term, but given that both are sortal terms this is quite irrelevant to the points in dispute between myself and Lowe.

University of Birmingham,
P.O. Box 363, Birmingham B15 2TT

© HAROLD W. NOONAN 1986

SUPERMAN, THE IMAGE

By ANDY CLARK

THERE has been much spurious debate over the issue of mental imagery (see Dennett [1] for a decent burial). There is, however, a genuine question to be addressed, namely, is the capacity to form a mental image dependent on some special kind of internal representation[1] – a kind of representation formally distinct from that which underlies our non-imagistic capacities of recall and thought? Those who favour a positive answer to this question usually stress some kind of structural isomorphism between the mental image and the object it represents. This claim may be put by saying that the representation is analogical (e.g. Metzler and Shepherd [2]), or that it depends on data in a 2D array (Kosslyn *et al.* [3]), or that it is somehow pictorial or holographic or any one of a host of other locutions. The central point, at any rate, is that having a mental image is said to involve cognitive operations closely akin to those involved in *ordinary perception.* I propose a simple experiment to cast doubt on this claim, and hence on the claim that having mental images involves deploying some special kind of underlying representation.

Imagine (*sic*) the following situation. I am sitting with a pen in my hand and a blank sheet of paper in front of me. Daydreaming, I imagine Superman flying through the air. I can even, it seems, locate my image in the centre of the blank sheet of paper. It occurs to me to try to (as it were) trace the image. That is, I try to trace a picture of Superman by going around the edges of the projected image with the pen. But the project is a dismal failure. The puzzle then is how such a failure is possible. It need not be inevitable – the puzzle is genuine just so long as it is possible for even one person to be capable of visualizing the image yet incapable of tracing it. I am certainly such a person, and I expect there are many others.

How might we explain this phenomenon? The following possibilities have been suggested[2].

(a) A simple failure of motor control is to blame; my hand-eye co-ordination is simply inadequate. – This won't do, as I am quite capable of tracing a *real* picture of Superman.

(b) The image is in 3D and hence (in some rather obscure way) the wrong shape to trace. – Insofar as I understand this, it can't be right. My image of Superman is derived from exposure to comics and films. Why shouldn't the image likewise be classed as a 2D representation of a 3D object?

---

[1] This formulation is given by P. N. Johnson-Laird in his commentary on Kosslyn *et al* [3] (reference as [3] p. 557-8).

[2] By members of the Cognitive Studies Research Programme at Sussex. Thanks especially to Dave Young and Tom Khabaza.

(c) The image won't stablize sufficiently to be traced. – This seems more promising. The idea is that my image may keep subtly shifting between different poses thereby disorienting me in my attempts to trace it. But it is a fairly simple matter to induce stability. Just imagine that Superman is holding up a falling wall. This fixes the pose, but all attempts to trace it still fail.

(d) A 'zoom effect' is systematically de-stabilizing the image. – The thought here is that as soon as I try to trace some specific part of the image (say, Superman's foot) then that part of the image 'reacts' by seeming to expand and thus disturbing my grasp of its relation to the next part I want to trace. Here I think we may have something important. But the question to push now is *why* should such a 'zoom effect' occur? After all, no such problem plagues us when we try to trace a *real* picture.

I propose that the correct explanation of both the tracing problem *and* the (putative) zoom effect is that the information which we introspectively treat as an image is *not*, in fact, in any kind of analogical, pictorial or otherwise 'structurally isomorphic' form at all. This is *not* to say it is in the form of propositions, and *certainly* not propositions in our natural language. We can, after all, imagine much that we cannot fully describe. Rather, I take the tracing problem phenomenon to indicate merely that *whatever* mode of representation underlies our image-forming talents, it does not involve anything closely akin to the cognitive operations by which we perceive a real-world scene. The cognitive operations and representations may instead be more like those employed in, say, recalling a telephone conversation, or the results of an experiment. The zoom effect could then be explained as simply the calling up of more data concerning whatever we have just focused our attention on. (Just as I may focus first on the idea of Superman's weaknesses and then find myself accessing data about kryptonite and red suns.) If the process of consulting an image is not a 'quasi-perceptual' one, then the zoom effect and the tracing problem need come as no surprise.

Against this contention it may be urged that other results strongly suggest that we solve some problems by manipulating pictorial representations. Thus Metzler and Shepherd [2] asked subjects whether a picture depicts an object which could be rotated to look like a second picture which they are shown. The subjects answered in a time proportional to the amount of rotation needed to turn the object so as to present the target view. The obvious explanation is that the subjects answered by literally rotating a mental image.

Such results are by no means conclusive. Our question , after all, is in what underlying mode of representation this 'rotation' was accomplished, and the result does not rule out some regular systematic but non-analog process. Moreover, other results (Rock [4]) establish that our capacities of rotation vary with the complexity of the figure, and that subjects are sometimes unable to answer basic questions concerning the final view of a mentally rotated image (e.g. they may be unable to say how many corners of a cube are 'visible' after certain rotations – see Hinton [5]). 'Mental rotation' evidence is thus at best ambiguous, and I am inclined to agree with Pylyshyn [6] that it provides insufficient evidence for 'jumping on the analogical bandwagon'.

It would be rash to conclude that a dualism of types of internal representation is out of the question. But the tracing problem strongly suggests that if images *are* special, it is *not* because they involve quasi-perceptual cognitive processes. In short, there may be less to mental images than seems to meet the eye.

University of Sussex,
School of Social Sciences,
Arts Building, Falmer,
Brighton BN1 9QN

© ANDY CLARK 1986

REFERENCES

[1] D. Dennett, 'Two approaches to mental images' in *Brainstorms*, Dennett, D. (Harvester Press, Sussex, 1981).
[2] J. Metzler, and R. N. Shepherd, 'Transformational studies of the internal representation of three-dimensional objects' in R. L. Solso (ed.) *Theories of Cognitive Psychology, the Loyola Symposium* (Lawrence Erlbaum, Hillsdale N. J., 1974).
[3] S. M. Kosslyn, S. Pinker, G. Smith, and S. P. Schwartz, 'On the demystification of mental imagery' *Behavioural and Brain Sciences* 2:4, 1979 p. 535-48.
[4] J. Rock, *Orientation and Form* (Academic Press, New York, 1973.)
[5] G. Hinton, 'Some demonstrations of the effects of structural descriptions in mental imagery', *Cognitive Science* 3 (3) p. 231-51, 1979.
[6] Z. Pylyshyn, 'Imagery and artificial intelligence' in *Minnesota Studies in the Philosophy of Science*, ed. C. Wade Savage, Vol. IX (University of Minnesota Press, 1978).

CORRECTION

In Anne Jaap Jacobson's article 'Causality and the Supposed Counterfactual Conditional in Hume's *Enquiry*' in ANALYSIS 46.3, June 1986, the following corrections should be made:

p. 131, line 9, insert '(a component of)' between 'that' and 'the causal relationship';

p. 131, line 10, delete '[4] p. 173' and substitute '[1] p. 174 and fn. 18'.

CALL FOR PAPERS

ARGUMENTATION

An International Journal on Reasoning

First Issue: Spring 1987

D. Reidel Publishing Company are pleased to announce that in the Spring of 1987 they will commence publication of *Argumentation*, an International Journal on Reasoning.

Scope. Argumentation has seen a major revival in the twentieth century, in the United States as well as in Europe, under the influence of the philosophy and linguistic analysis of everyday reasoning. Interest in rhetoric has now extended beyond the traditional boundaries set by the classics.

Argumentation is a new international journal. Its aim is to gather contributions from all schools of thought – ranging from literary rhetoric to linguistics, from history to logic, from theological arguments to legal reasoning, from natural inference to the argumentative structures of science.

This journal is of special concern to specialists in speech and communication, philosophers, literary critics, and linguists, as well as all those interested in natural and artificial intelligence.

The Editors are Jean-Blaise Grize (Neuchâtel, Editor-in-Chief), Michel Meyer (Brussels), Frans van Eemeren (Amsterdam). Managing Editors: Christian Plantin (Brussels), Rob Grootendorst (Amsterdam). Editorial Board: Jean-Claude Anscombre (Paris), Marie-Jeanne Borel (Lausanne), Marc Dominicy (Brussels).

Manuscripts for publication (the original + two copies) should be sent to Dr Michel Meyer, European Centre for the Study of Argumentation (CEEA), Université Libre de Bruxelles, Institut de Philosophie, 143 avenue A. Buyl, C. P. 188, 1050 Brussels, Belgium. The author should retain an additional copy of the manuscript for reference. Unless a specific request is made, copies will not be returned.

The principal language of the journal is English, but papers in French are also accepted.

Enquiries concerning Information for Authors, sample copies, and subscriptions should be addressed in the first instance to: The Journals Department, D. Reidel Publishing Company, P.O. Box 17, 3300 AA Dordrecht, Holland. (Telephone: (31)78–334208. Telex: 29245 reid nl.)

the review of

metaphysics

a philosophical quarterly

JUNE 1986 | VOL. XXXIX, No. 4 | ISSUE No. 156 | $8.00

articles

E. M. ADAMS	*The Human Substance*
JOSEPH MARGOLIS	*Constraints on the Metaphysics of Culture*
KENNETH L. SCHMITZ	*Metaphysics: Radical, Comprehensive, Determinate Discourse*
M. M. SCHUSTER	*Is the Flow of Time Subjective?*
A. SHALOM	*Culture and Psychoanalysis*
DOROTHEA FREDE	The Impossibility of Perfection: *Socrates' criticism of Simonides' Poem in the* Protagoras
JANICE FLOERSCH and Staff	*Summaries and Comments*

philosophical abstracts *announcements*

Individual Subscriptions $21.00 Institutional Subscription $32.00 Student Subscriptions $1

Jude P. Dougherty, Editor
The Catholic University of America, Washington, D.C. 20064

MIND

Editor: Simon Blackburn, Pembroke College, Oxford

Mind is one of the most widely read and respected journals in professional philosophy.

Forthcoming and recent papers include *George Boolos* on Frege, *Michael Tye* on the Subjective Character of Experience, *André Gallois* on Rigid Designation and the Contingency of Identity, *Jennifer Hornsby* on Non Indicatives in Truth Theories, and *Susan Hurley* on Supervenience.

Rates for Volume 95, 1986: UK £10 N. America $22.50 Elsewhere £13.50
Institutions: UK £16 N. America $32.50 Elsewhere £19

ORDER FORM

☐ Please enter my subscription to **Mind** Vol. 95 (1986)

☐ Please send me a sample copy

☐ I enclose my correct remittance (payment may be made by personal cheque, credit transfer or major credit cards.)

Name ....

Address ....

Postcode........................ Country .... XA

OXFORD JOURNALS

Journals Subscription Department, OUP, Walton Street, Oxford OX2 6DP, UK

NOTES

The ANALYSIS Committee consists of: Chairman, P. T. Geach; Secret T. R. Baldwin; T. S. Champlin, M. Clark, J. P. Dancy, Andrew Harrison, Heal, R. F. Holland, J. H. McDowell, D. H. Mellor, H. W. Noonan, P. Sm M. A. Stewart, R. G. Swinburne, A. R. White, C. J. F. Williams, Peter Wi This committee is responsible for appointing and advising the Editor and the general policy of the paper.

SUBSCRIPTIONS. ANALYSIS is published by Basil Blackwell Ltd, 108 Cow Road, Oxford OX4 1JF, England and P.O. Box 1320, Murray Hill Stat NY 10156, U.S.A. Each volume comprises four numbers, each of at least pages, appearing in January, March, June and October. Orders should be to the publisher. Subscriptions prices for 1986 are as follows:

	U.K.	Overseas (incl. Eire)	U.S.A. & Canada
Institutions	£11.75	£14.50	U.S. $33.00
Individuals	£7.85	£9.75	U.S. $21.00

BACK ISSUES. Volumes 1–37 are now available in complete volumes in offset reprint from Swetz & Zeitlinger, 471 Keizersgracht, Amsterd Holland. Recent back issues are available from the publisher at the cur single issue price.

MICROFILM copies of complete volumes are available from Volume 1 wards from University Microfilms International Ltd, 30–32 Mortimer Str London W1N 7RA, England, or 300 North Zeeb Road, Ann Arbor, Mich 48106, U.S.A.

ADVERTISING. For details contact Bill Drake, telephone (0869) 38477 write c/o the publishers.

CONTRIBUTIONS. Articles submitted for publication should be addre to Mr Christopher Kirwan, Exeter College, Oxford OX1 3DP.

Contributors are asked to note the following requirements. Articles sho normally not exceed 3,000 words in length, including references and footnot They should be type written in double spacing on one side of the paper on photographic copies are acceptable. Footnotes should be kept to a minim and wherever possible avoided altogether. Inset quotations and form should be well spaced from the text preceding and following them. The spac of symbols within formulae should be carefully attended to. Single quotati marks should be used, except for purposes of internal quotation and "sca quotes. Discussion papers should be sent in as soon as possible after appearance of the article to which they refer.

Typescripts of rejected articles will be returned only if self-addre envelopes with correct postage are provided (in the case of overseas contribu international reply coupons sufficient for airmail).

Proofs of accepted articles will be sent to authors for correction, toget with information about free copies. Typescripts of accepted articles will retained by the Editor on the assumption that authors have kept their copies.

Apart from correction of printers' errors, only very minor changes in p will normally be accepted by the Editor, to avoid delay and expense production.

Typeset by Unicus Graphics Ltd, Horsham, Printed and bound in Great Britain
Whitstable Litho Ltd, Whitstable